

UNIVERSITÉ DE SHERBROOKE

Faculté d'éducation

Utilisation d'un instrument d'évaluation authentique (EIS) afin de corroborer l'accès aux services  
chez les jeunes enfants

par

Maxime Poitras

Mémoire présenté à la Faculté d'éducation

en vue de l'obtention du grade de

Maître ès Science (M. Sc.)

Maîtrise en psychoéducation

Septembre 2020

© Maxime Poitras, 2020

UNIVERSITÉ DE SHERBROOKE

Faculté d'éducation

Utilisation d'un instrument d'évaluation authentique (EIS) afin de corroborer l'accès aux services  
chez les jeunes enfants

par

Maxime Poitras

Ce mémoire a été évalué par un jury composé des personnes suivantes :

Malena Argumedes Charles  
Université de Sherbrooke

Directrice ou directeur de la recherche

Carmen Dionne  
Université du Québec à Trois-Rivières

Codirectrice ou codirecteur de la recherche

Alexa Martin-Storey  
Université de Sherbrooke

Membre du jury

Mémoire accepté le 13 novembre 2020

## **SOMMAIRE**

Le retard global de développement (RGD) est un diagnostic souvent temporaire utilisé auprès des jeunes enfants lorsque qu'ils ne sont pas en mesure d'être soumis aux évaluations standards pour diverses raisons comme leur âge ou la nature de leurs difficultés. Selon le Ministère de la Santé et des Services Sociaux du Québec le RGD se définit comme étant un retard, chez un enfant âgé de moins de sept ans, équivalent ou supérieur à deux écart-types sous la moyenne établie pour son âge chronologique dans au moins deux des sphères développementales suivantes : motricité, cognition, socioémotionnelle, communication et autonomie. Il s'agit d'une condition qui touche entre 1 % et 3 % des enfants, soit l'équivalent de la déficience intellectuelle. Il s'agit également de la problématique la plus rapportée chez les enfants âgés entre 0 et 4 ans. Les causes du RGD sont multiples et demeurent très mal comprises par les chercheurs et les cliniciens. Il s'agirait en fait d'un ensemble de conditions qui s'additionnent et en viennent à nuire au développement de l'enfant de manière significative. Les conséquences sont également très hétérogènes. La plupart des enfants recevant un diagnostic de RGD seront ultérieurement diagnostiqués comme présentant une déficience intellectuelle. De plus, on note des problématiques au niveau de la régulation émotionnelle, des fonctions exécutives, des habiletés sociales et des capacités adaptatives.

Au fil des ans, plusieurs instruments d'évaluation du développement ont été rendus disponibles pour les intervenants œuvrant auprès de cette clientèle. La majorité de ces instruments sont standardisés et ont été créés dans le but de déterminer un âge de développement pour l'enfant. Bien que plusieurs de ces instruments possèdent des caractéristiques psychométriques

satisfaisantes appuyées par plusieurs études, plusieurs critiques en lien avec l'utilisation de ce type d'instrument auprès de la clientèle présentant un RGD sont présentes dans la littérature. Ces critiques touchent la fidélité et la validité des instruments standardisés lorsqu'ils sont utilisés spécifiquement auprès des enfants présentant un RGD ainsi que l'interprétation de leurs résultats. C'est dans une tentative de palier les différentes lacunes des instruments standards que l'évaluation authentique se développe. Il s'agit d'un type d'évaluation, pouvant être utilisé en complémentarité avec les évaluations standards et qui ne vise donc pas à les remplacer, axée sur l'observation et les activités quotidiennes et dont la passation s'effectue dans les milieux naturels des enfants. Concrètement, l'évaluation authentique se concentre sur ce que l'enfant est en mesure de faire compte tenu de son niveau de développement et non sur l'obtention d'un score lié à un âge développemental. Toutefois, très peu d'instruments d'évaluation authentique dont les caractéristiques psychométriques ont été documentées sont disponibles en français.

Le programme EIS est un instrument d'évaluation authentique se basant sur un curriculum. L'outil permet de dresser un portrait global et organisé de l'ensemble du développement et des comportements de l'enfant tout en facilitant la transition entre l'évaluation et l'intervention. Dans la version anglophone de l'EIS, des scores seuils ont été développés afin de permettre aux intervenants d'utiliser l'EIS pour corroborer l'admissibilité aux services d'intervention précoce spécialisés d'enfants présentant potentiellement un RGD.

Le présent mémoire vise donc à contribuer au développement de la littérature francophone sur les évaluations authentiques du développement en petite enfance en comparant les résultats obtenus par 21 enfants à l'EIS avec ceux obtenus à l'*Ages and Stages Questionnaires*, 3<sup>e</sup> édition, un

instrument de dépistage possédant de bonnes caractéristiques psychométriques. Des corrélations de Pearson ont été utilisées pour comparer les deux instruments. Les résultats montrent que les corrélations ne sont pas significatives. Toutefois, il est possible de constater que la grande majorité des résultats obtenus pour les 21 enfants sont identiques à l'EIS et à l'ASQ-3, c'est-à-dire qu'ils se trouvent au-dessus du score seuil. Au total, sept résultats divergent en étant sous le point de coupure à l'EIS alors que cela n'était pas le cas à l'ASQ-3.

Ces résultats laissent croire que l'EIS est en mesure d'identifier les enfants ne présentant pas de RGD avec un certain niveau de précision dans les domaines cognitif et de la motricité fine. De plus, il s'avère que les résultats divergents concernent principalement une tranche d'âge en particulier et que l'ampleur de cette divergence est très faible pour la motricité fine. Toutefois, elle est plus élevée pour le domaine cognitif. Plusieurs limites liées au recrutement ainsi qu'à l'échantillon pourraient expliquer ces résultats. Des études futures utilisant un échantillon plus important ainsi que des enfants connus pour présenter un RGD devraient être réalisées. Néanmoins, les résultats de ce projet ajoutent à la littérature francophone concernant le développement d'instruments d'évaluation authentique francophones et pourraient être utilisés pour jeter les bases à de nouveaux projets.

## TABLE DES MATIÈRES

<b>INTRODUCTION.....</b>	<b>1</b>
<b>LA PROBLÉMATIQUE .....</b>	<b>2</b>
<b>PREMIER CHAPITRE.</b>	
1. LE RETARD GLOBAL DE DÉVELOPPEMENT.....	4
1.1 Définitions.....	4
1.2 Prévalence, conséquences et étiologie .....	6
2. L'ÉVALUATION DU DÉVELOPPEMENT DE L'ENFANT .....	8
2.1 L'évaluation à l'aide d'outils standardisés du développement.....	9
3. ÉLIGIBILITÉ AUX SERVICES D'INTERVENTION PRÉCOCES .....	11
3.1 L'accès aux services spécialisés aux États-Unis et au Québec .....	11
3.2 Les services spécialisés d'intervention précoce québécois .....	13
4. ÉVALUATION AUTHENTIQUE.....	13
4.1 Définition de l'évaluation authentique .....	14
4.2 Les évaluations critériées du développement.....	15
4.3 Évaluation basée sur des curriculums .....	16
4.4 Questions de recension.....	17
<b>LA RECENSION DES ÉCRITS.....</b>	<b>19</b>
<b>DEUXIÈME CHAPITRE.</b>	
1. MÉTHODE DE RECENSION .....	19
2. CARACTÉRISTIQUES DES ÉTUDES RECENSÉS .....	21
2.1 Études empiriques .....	21
2.2 Autres études sur le sujet .....	24

<b>3.</b>	<b>LES RÉSULTATS DES ÉTUDES .....</b>	<b>27</b>
3.1	Objectif 1 : les instruments développés .....	27
3.2	Objectif 2 : les caractéristiques psychométriques des instruments .....	29
3.3	Objectif 3 : Pertinence des évaluations authentiques .....	31
3.4	Les recherches futures proposées.....	34
3.5	Synthèses des études recensées.....	35
<b>4.</b>	<b>RETOMBÉES POUR LA PSYCHOÉDUCATION .....</b>	<b>38</b>
<b>5.</b>	<b>OBJECTIFS INITIAUX POURSUIVIS .....</b>	<b>40</b>
<b>6.</b>	<b>CONTEXTE PARTICULIER DU MÉMOIRE .....</b>	<b>40</b>
	<b>LA MÉTHODE .....</b>	<b>43</b>
	<b>TROISIÈME CHAPITRE.</b>	
<b>1.</b>	<b>CONTEXTE DE L'ÉTUDE .....</b>	<b>43</b>
<b>2.</b>	<b>LE DEVIS DE L'ÉTUDE.....</b>	<b>43</b>
<b>3.</b>	<b>PARTICIPANTS .....</b>	<b>43</b>
3.1	Sélection des enfants non éligibles aux services.....	44
<b>4.</b>	<b>INSTRUMENTS DE MESURE.....</b>	<b>44</b>
4.1	Questionnaire sociodémographique.....	45
4.2	Le Age & Stages Questionnaires 3 (Squires et Bricker, 2009).....	45
4.3	Le Programme Évaluation, Intervention, Suivi (Johnson et Bricker, 2006) .....	47
4.3.1	Caractéristiques psychométriques de l'EIS.....	49
<b>5.</b>	<b>DÉROULEMENT .....</b>	<b>51</b>
<b>6.</b>	<b>MÉTHODE D'ANALYSE DE DONNÉES .....</b>	<b>52</b>
	<b>LES RÉSULTATS .....</b>	<b>53</b>
	<b>QUATRIÈME CHAPITRE.</b>	
<b>1.</b>	<b>STATISTIQUES DESCRIPTIVES DES PARTICIPANTS.....</b>	<b>53</b>

2.	COMPARAISON ENTRE LES INSTRUMENTS .....	54
	<b>LA DISCUSSION.....</b>	<b>58CINQUIÈME CHAPITRE.</b>
1.	RAPPEL DE L'OBJECTIF DU PROJET.....	58
2.	COMPARAISON ENTRE LES DEUX INSTRUMENTS .....	58
3.	LIMITES DU PROJET .....	61
4.	RECHERCHES FUTURES .....	62
	<b>CONCLUSION .....</b>	<b>63</b>
	<b>RÉFÉRENCES BIBLIOGRAPHIQUES.....</b>	<b>65</b>



## **LISTE DES TABLEAUX**

Tableau 1.	Caractéristique des études empiriques.....	22
Tableau 2.	Résultats au domaine motricité fine .....	53
Tableau 3.	Résultats au domaine cognitif... ..	54

## **LISTE DES ABRÉVIATIONS, DES SIGLES ET DES ACRONYMES**

ASQ-3	Ages and Stages Questionnaires, 3 <sup>e</sup> édition
CIUSSS	Centre intégré universitaire de santé et services sociaux
CISSS	Centre intégré de santé et services sociaux
CPE	Centre de la petite enfance
DEC	Division for Early Childhood
ÉA	Évaluation authentique
ÉBC	Évaluation basée sur un curriculum
EIS	Le programme Évaluation, Intervention et Suivi
IDEA	Individuals with Disabilities Education Act
INSPQ	Institut national de santé publique du Québec
MSSS	Ministère de la Santé et des Services Sociaux
OPPQ	Ordre des psychoéducateurs et psychoéducatrices du Québec
RGD	Retard global de développement
TSA	Trouble du spectre de l'autisme

## REMERCIEMENTS

Je désire tout d'abord remercier ma directrice ainsi que ma codirectrice, soit les professeures Malena Argumedes et Carmen Dionne pour leur inestimable soutien tout au long de cette longue aventure. Votre passion, votre leadership et votre grande compétence m'ont permis de demeurer motivé dans les moments plus difficiles et m'ont poussé à vouloir être un peu meilleur chaque jour. La richesse de votre support me permet aujourd'hui de ressortir grandi de cette expérience, tant sur le plan professionnel que personnel. Des remerciements sont également adressés au professeur Michel Rousseau pour sa disponibilité et son généreux support.

Un remerciement particulier à la professeure Colombe Lemire qui, tantôt par une blague, tantôt par une discussion hors sujet et souvent par un conseil attentionné semble toujours être en mesure de dire les bonnes choses au bon moment.

Je ne peux passer sous silence la contribution de l'équipe de la Chaire UNESCO dépistage et évaluation du développement des jeunes enfants avec qui je travaille depuis plus de cinq ans. Merci aux professeures Carmen Dionne et Annie Paquet pour les nombreuses opportunités m'ayant permis de développer une passion pour la recherche. Un sincère remerciement à Maude Boutet, professionnelle de recherche, pour la compréhension et la bienveillance dont tu as fait part ces dernières années. Merci à Annie Claude Dubé pour ton implication dans le recrutement des participants. Je me dois également de remercier Marie-Joëlle, Annie-Claude Bourque, Christina, Vicky et Roxanne avec qui j'ai partagé de beaux moments comme assistant de recherche au travers les années.

Un remerciement bien spécial à mes parents Dominique et Claude pour votre inconditionnel soutien et les sacrifices réalisés depuis près de 25 ans. Vous avez toujours été un modèle de persévérance et de travail acharné. Merci de m’avoir encouragé à foncer tête baissée dans ce qui me passionne, ce sentiment de confiance fut une source infinie d’énergie. Je dois également remercier Lauriane pour ta compréhension, ton écoute et ton précieux soutien lors de cette dernière année rocambolesque. Un petit clin d’œil est dirigé vers Raphaël, Simon et Alexis, des amis qui, probablement sans s’en rendre compte, ont joué un rôle important dans l’accomplissement de ce projet.

Je désire également remercier les CPE ayant participé au projet. Votre disponibilité et votre engouement pour la recherche m’a permis de matérialiser l’ensemble du projet. Sans votre précieuse collaboration et vos efforts, rien de tout cela ne serait possible. Merci de contribuer au développement des connaissances concernant l’évaluation des jeunes enfants.

Ce mémoire a reçu un soutien financier de la part de l’Équipe des Fonds de recherche du Québec — Société et culture — Pour des leviers multiples d’intervention (Équipe FRQSC ; 2019, 2020), ainsi que du Groupe de recherche sur l’intervention précoce inclusive (GRIPI ; 2019, 2020). Des bourses d’aide à la diffusion pour des communications scientifiques ont également été obtenues par l’Université du Québec de Trois-Rivières (UQTR ; 2019, 2020), l’Équipe FRQSC — Pour des leviers multiples d’intervention (2019, 2020) et le GRIPI (2019, 2020). Merci à ces différents organismes pour leur soutien.

## INTRODUCTION

Le retard global de développement (RGD) est un diagnostic dont nous entendons moins souvent parler. Toutefois, il s'agit d'un diagnostic qui touche plusieurs enfants au Québec et ailleurs dans le monde et qui bouleverse souvent la réalité quotidienne des familles. Différentes évaluations sont disponibles pour permettre aux professionnels de détecter et d'agir rapidement afin de limiter l'impact que le RGD peut avoir. Les interventions subséquentes sont directement en lien avec la qualité et l'évaluation réalisée au départ puisque celle-ci permet de colliger un nombre élevé d'informations précieuses pour l'équipe d'intervenants.

Ce mémoire s'intéresse d'ailleurs aux évaluations du développement des jeunes enfants et plus particulièrement ceux présentant ou étant à risque de présenter un RGD. Le premier chapitre présente la problématique en définissant le RGD ainsi que sa prévalence, ses causes et ses conséquences. L'évaluation du développement de l'enfant est ensuite abordée de même que les deux types d'instruments souvent utilisés, soit les évaluations standardisées du développement et les évaluations dites authentiques du développement de l'enfant. Les mécanismes d'accès aux services d'intervention spécialisés sont ensuite décrits. Une recension des écrits constitue le deuxième chapitre et se centre sur les différents instruments disponibles dans la littérature et la détermination de l'objectif final du projet. Au chapitre trois, la méthode vient ensuite présenter la procédure utilisée, les participants au projet ainsi que les instruments utilisés. Les résultats sont abordés au chapitre quatre avant d'être discutés au chapitre suivant. La discussion clôt le mémoire en abordant la portée des résultats, les limites de l'étude ainsi que les prochaines étapes à réaliser.

## **PREMIER CHAPITRE. LA PROBLÉMATIQUE**

Ce projet de recherche vise à explorer la possibilité d'utiliser les résultats d'un instrument d'évaluation authentique afin de corroborer l'éligibilité aux services spécialisés d'intervention précoce pour les enfants de moins de six ans présentant un retard global de développement (RGD). Cette évaluation authentique est le test de la deuxième édition du Programme Évaluation, Intervention et Suivi (Bricker, 2006). Le programme Évaluation, Intervention et Suivi (EIS) est un instrument utilisé pour la planification et la mise en œuvre de l'intervention auprès des jeunes enfants. Dans le cadre de ce mémoire, des analyses statistiques seront utilisées afin de déterminer si les points de coupures de deux domaines (motricité fine et cognitif) disponibles dans la version américaine de l'instrument peuvent identifier efficacement les enfants ayant besoin de services au Québec. Ces points de coupure permettent de déterminer si le développement de l'enfant est significativement inférieur à la moyenne dans ces domaines. Ce travail s'inscrit dans une amorce de changement paradigmatique en ce qui a trait à l'évaluation du RGD amorcée au début des années 2000 aux États-Unis (McLean, 2005). Une remise en question quant aux méthodes d'évaluation servant à juger de l'éligibilité aux services des jeunes enfants émerge de recherches soulignant les limites des tests standardisés actuellement utilisés à cet effet (Macy, M., Bagnato, Macy, R. et Salaway, 2015). Différentes limites liées à ces évaluations ont été soulignées, notamment en ce qui a trait à leur validité. Celles-ci seront abordées plus loin. Au Québec, l'Ordre des Psychoéducateurs et Psychoéducatrices du Québec (OPPQ; 2013) recommande l'utilisation de tests standardisés et fonctionnels, c'est-à-dire basés sur les activités de la vie quotidienne, dans l'évaluation du RGD. Malgré cela, les évaluations standardisées du développement de l'enfant

demeurent la seule méthode d'évaluation utilisée dans la majorité des cas (Macy et Bagnato, 2010; Rousseau, Dionne et Paquet, 2012). Afin de pallier les différentes limites de ces tests, des stratégies alternatives ont été développées ces dernières années dans le domaine de l'intervention précoce. Parmi elles se trouve l'évaluation authentique (ÉA) qui regroupe diverses méthodes d'évaluation, notamment les évaluations du développement de l'enfant basées sur un curriculum dont fait partie l'EIS. Il s'agit d'une pratique d'évaluation recommandée par le National Association for Education of Young Children (NAEYC) (2003) depuis plus de dix ans.

Le projet original tel qu'il avait été développé et réfléchi a dû faire l'objet de modifications importantes en raison des mesures sanitaires mises en place suite à la pandémie de COVID-19. Ces changements seront expliqués en détails dans la section objectifs.

Ce premier chapitre est destiné à la présentation de la problématique du projet de recherche. Dans un premier temps, le retard global de développement sera défini. Les conséquences connues du RGD ainsi que les pratiques d'évaluation seront ensuite présentées. Le portrait des procédures d'évaluation utilisées pour déterminer l'éligibilité aux services d'intervention précoce sera détaillé pour les États-Unis et le Québec. Les critiques effectuées par la communauté scientifique à l'égard des évaluations standardisées du développement de l'enfant seront également discutées. Finalement, un bref historique de l'ÉA ainsi que les différentes définitions utilisées dans la littérature seront présentés.

## 1. LE RETARD GLOBAL DE DÉVELOPPEMENT

Afin de bien cerner le concept de retard global de développement, l'utilisation qui est faite de ce diagnostic dans les milieux cliniques doit être expliquée. Dans un premier temps, il importe de préciser que le RGD est généralement utilisé en tant que diagnostic temporaire auprès des jeunes enfants, ces derniers n'étant pas nécessairement en mesure de se soumettre aux évaluations standards pour diverses raisons, notamment le fait qu'ils n'ont pas atteint l'âge minimum pour la passation de certains tests (American Psychiatric Association, 2013). À cet effet, le *Manuel diagnostique et statistique des troubles mentaux, 5<sup>e</sup> édition* (DSM-5) mentionne qu'il est nécessaire de réévaluer de nouveau un enfant dans le futur lorsque ce diagnostic lui est attribué. De plus, dans le présent mémoire, le RGD sera utilisé pour décrire un retard significatif dans au moins deux sphères de développement de l'enfant. Les retards non significatifs ou les retards de développement ne touchant qu'un seul domaine de développement ne sont donc pas considérés.

### 1.1 Définitions

La définition la plus utilisée en ce qui a trait au RGD est celle inscrite au DSM-5 (APA, 2013) et est libellée comme suit :

Ce diagnostic est réservé aux enfants de moins de 5 ans, lorsque le niveau de sévérité clinique ne peut être établi en raison d'un manque de fidélité dans l'évaluation et l'utilisation d'instruments standardisés. Ce diagnostic est posé lorsque l'enfant n'évolue pas selon les stades de développement attendus, et ce, dans plusieurs domaines de fonctionnement. Ce diagnostic s'applique également aux enfants dont il est impossible d'évaluer le fonctionnement intellectuel de manière systématique, c'est-à-dire les enfants trop jeunes pour qu'on utilise avec eux des outils



standardisés. Cette catégorie exige également une réévaluation après un laps de temps. (p. 41)

Au Québec, le Ministère de la Santé et des Services Sociaux (MSSS; 2017) emploie l'appellation « Retard significatif de développement » pour faire référence aux enfants de moins de sept ans ayant un retard équivalent à deux écarts-types sous la moyenne établie pour leur âge chronologique dans au moins deux des sphères de développement suivantes : motricité, cognition, socioémotionnelle, communication et autonomie. L'évaluation doit avoir été faite à l'aide d'un outil normalisé. Celle-ci vise à situer l'enfant comparativement aux autres enfants de son âge en ce qui a trait à son développement (OPPQ, 2014). Cette définition englobe également les enfants ayant reçu un diagnostic de RGD tel que prévu au DSM-5 ainsi que ceux en attente d'une évaluation ou dont le retard significatif n'est explicable par aucun autre trouble précis. De son côté, l'OPPQ (2013) utilise la définition de Shevell *et al.* (2003) dans son document dictant les lignes directrices liées à l'évaluation du trouble. Notons toutefois que ces lignes directrices sont présentement en réédition. Cette définition réfère à un retard significatif dans deux domaines ou plus entre les suivants : motricité fine et globale, parole et langage, cognitif, social et personnel et activités de la vie quotidienne chez un enfant de cinq ans ou moins et correspondant à deux écarts-types sous la moyenne. Comme pour le DSM-5, l'évaluation doit avoir été réalisée à l'aide d'un outil normalisé.

Dans le présent projet de mémoire, le terme *retard global de développement* réfère à la définition utilisée par le MSSS (2017). Le choix de cette définition découle de trois raisons. Premièrement, elle permet d'inclure les enfants ayant été évalués pour un RGD par le biais

d'évaluations normatives ciblant des domaines développementaux précis. Deuxièmement, elle inclut la définition du DSM-5 (APA, 2013). Finalement, cette définition sera utilisée puisqu'elle provient du cadre de référence gouvernemental visant l'organisation des services dispensés aux jeunes ayant un RGD dans la province de Québec et que les participants de la présente étude reçoivent des services de cette organisation.

## **1.2 Prévalence, conséquences et étiologie**

Peu de données fiables sont en mesure de statuer de la prévalence du RGD. En 2003, Shevell *et al.* affirmaient qu'entre 1 % et 3 % des enfants présentaient un RGD, ce qui est l'équivalent de la prévalence de la déficience intellectuelle. Au Canada, l'Enquête sur la participation et les limitations d'activités (Statistique Canada, 2001), utilisant des données autorapportées, suggère que le RGD serait l'incapacité la plus présente chez les jeunes enfants, représentant 68 % des incapacités chez les 0 à 4 ans et 1,1 % de la population d'enfants en général. Dans une étude publiée en 2006, Srour, Mazer et Shevell se sont attardés à l'étiologie du RGD. Leurs résultats montrent que les causes du RGD sont souvent multiples et majoritairement inconnues.

Les conséquences du RGD sont très hétérogènes et une multitude de facteurs distaux et proximaux peuvent les influencer. Cela est dû au fait que seulement certaines sphères de développement peuvent être affectées chez l'enfant. Ainsi, un jeune dont les domaines cognitif, de la communication et socioaffectif sont déficitaires aura un profil fort différent d'un autre pour qui les domaines touchés sont la communication et les capacités adaptatives. Bien qu'il ne soit pas

impératif qu'un retard dans le domaine cognitif ou intellectuel soit présent pour recevoir un diagnostic de RGD, la vaste majorité des enfants recevant un diagnostic de RGD en recevront un de déficience intellectuelle quelques années plus tard (Shevell, 2008). Cela s'explique par l'utilisation du RGD comme diagnostic temporaire en petite enfance auprès d'enfants qui seront réévalués dans le futur et pour lesquels il serait prématuré de conclure à des difficultés à ce niveau pour le moment. De manière générale, les enfants présentant un RGD ont un quotient intellectuel et verbal inférieur et atteignent certaines étapes développementales, comme apprendre à marcher, significativement plus tard que les enfants au développement typique (Kim, Jeon, Park, Chung et Song, 2014). Pour plusieurs des enfants touchés, les fonctions exécutives (mémoire de travail, planification et inhibition) présentent quelques faiblesses (Danielsson, Henry, Messer, et Rönnerberg, 2012). Ils ont également plus de difficultés au niveau de la régulation émotionnelle, des habiletés sociales et adaptatives ainsi qu'en résolution de problèmes sociaux et cognitifs (Baurain et Nader-Grosbois, 2009).

La recension des écrits de Bigras, Lemay et Tremblay (2012) traite des facteurs de risque et de protection associés au RGD. Dans cette étude, trois types de facteurs de risque sont identifiés comme ayant une grande incidence sur le développement potentiel d'un enfant. Premièrement, les facteurs personnels tels que le sexe masculin, les handicaps, un tempérament difficile et les maladies chroniques augmentent le risque de développer un RGD. Le deuxième type cible l'environnement familial. Ces facteurs comprennent, entre autres, la violence familiale, la dépression chez les parents et la qualité de la stimulation dispensée par le milieu familial.

Finalement, les facteurs environnementaux incluent la pauvreté, l'isolement ainsi qu'un environnement social de mauvaise qualité.

## 2. L'ÉVALUATION DU DÉVELOPPEMENT DE L'ENFANT

D'abord, il importe de préciser que le terme évaluation désigne, dans ce projet, les évaluations développementales de l'enfant. Bricker, Squires et Clifford (2010) proposent un modèle en quatre étapes correspondant au processus d'évaluation d'un enfant. Deux de ces étapes, soit la détermination de l'éligibilité aux services et le développement des buts et du plan d'intervention seront abordées. La détermination de l'éligibilité, parfois aussi appelée évaluation diagnostique, concerne les enfants identifiés comme étant à risque de développer des problèmes développementaux (Bricker *et al.*, 2010). Cette étape permet de juger, à l'aide d'instruments standardisés, si un jeune respecte les critères d'inclusion lui permettant de recevoir des services, c'est-à-dire les critères diagnostiques du RGD (Hallam, Lyons, Pretti-Frontczak et Grisham-Brown, 2014). L'étape suivante, destinée aux enfants jugés éligibles aux différents services, consiste en l'évaluation personnalisée et fonctionnelle de certains aspects spécifiques du développement de l'enfant dans l'optique d'identifier les cibles d'intervention pertinentes pour ce dernier (Bagnato, Neisworth et Pretti-Frontczak, 2010). Ces deux étapes font donc le pont entre l'entrée dans les services d'un jeune et le début de l'intervention. Bricker *et al.* (2010) affirment qu'il est possible d'associer ces deux étapes dans la mesure où une évaluation basée sur un curriculum est utilisée durant le processus de détermination de l'éligibilité d'un enfant. Ainsi, il serait possible d'utiliser un outil d'évaluation authentique basé sur un curriculum pour corroborer

ou déterminer l'éligibilité d'un enfant aux services (Bricker, Clifford, Yovanoff, Pretti-Frontczak, Waddell, Allen et Hoselton, 2008). Toutefois, l'utilisation d'outils d'évaluation authentiques visant à déterminer l'éligibilité d'un jeune enfant nécessitera de nombreux changements en termes de pratiques professionnelles (Bagnato *et al.*, 2010).

Afin de réaliser une évaluation juste et rigoureuse du développement d'un jeune, certains standards de qualité se doivent d'être respectés sans quoi, les résultats pourraient être invalides. Shevell (2008) souligne que l'évaluation du RGD devrait être multidimensionnelle, tant au niveau des professionnels (interdisciplinarité) que des modalités de cueillette de données (observation directe, évaluation par le jeu, etc.) et que les informations obtenues devraient provenir de différents milieux de vie et partenaires (parent, éducatrice, médecin, etc.). L'importance de diversifier les méthodes de cueillette de données semble également partagée par les intervenants effectuant les évaluations sur le terrain. Ces derniers considéreraient les résultats plus fidèles lorsque d'autres évaluations que celles traditionnellement utilisées leur permettent de corroborer leurs résultats (Keilty, LaRocco et Casell, 2009), particulièrement lorsque les méthodes utilisées documentent les habiletés fonctionnelles d'un jeune (Macy et Hoyt-Gonzales, 2007).

## **2.1 L'évaluation à l'aide d'outils standardisés du développement**

L'ensemble des définitions du RGD exigent que l'évaluation soit effectuée à l'aide d'outils standardisés, dits traditionnels, du développement de l'enfant. L'utilisation de ce type d'instrument est d'ailleurs commune et requise dans une majorité de pays occidentaux (Macy *et al.*, 2015). D'un point de vue théorique, ces tests reposent sur le principe que le développement d'un enfant se

distribue selon une courbe normale dans la population et qu'il est possible de situer un enfant par rapport à la moyenne des jeunes de son âge (McLean, 2005). Pour ce faire, des habiletés sont sélectionnées pour leur capacité à témoigner d'un âge de développement. Dans la littérature, bon nombre de ce type d'instruments sont disponibles pour les professionnels. Toutefois, certains questionnements concernant leur validité sont nommés (Bagnato, Goins, Pretti-Frontzcak et Neiswoth, 2014; Bagnato et Ho, 2006; Dalton et Brands, 2012; Freeman et Brown, 2008; Lazaro, 2017; Lee, Bagnato et Pretti-Frontzcak, 2015; Macy *et al.*, 2015; Macy et Bagnato, 2013; Neisworth et Bagnato, 2004).

Les critiques adressées envers les évaluations standardisées du développement de l'enfant, qui seront nommées évaluations traditionnelles, touchent différents aspects de leur utilisation. Dans un premier temps, de par leurs protocoles rigides de passation, elles ne permettent pas aux intervenants de procéder aux adaptations nécessaires à une passation valide en fonction des caractéristiques de l'enfant (Macy *et al.*, 2015). De plus, certains ne peuvent couvrir suffisamment de domaines de développement, le critère de référence étant de quatre domaines (Macy *et al.*, 2015). L'Institut National de Santé Publique du Québec (INSPQ, 2010) rappelle l'importance de documenter de nombreux domaines de développement. Le cas inverse a comme conséquence que l'enfant se voit dans l'obligation de subir plusieurs évaluations différentes (Macy *et al.*, 2015). Il importe également de mettre en évidence que ces évaluations dites traditionnelles du développement de l'enfant sont très peu en lien avec les habiletés fonctionnelles nécessaires au bon fonctionnement de l'enfant puisque les items ont été sélectionnés en fonction de leur capacité à témoigner d'un âge de développement et non en fonction de leur pertinence. De même, leurs

cueillettes de données s'effectuent rarement dans un contexte naturel (ex : dans une salle de rencontre inconnue de l'enfant; Bagnato et Ho, 2006; Kim, Lambert et Burts, 2013). Finalement, une fois l'éligibilité aux services établis, les intervenants se doivent de procéder de nouveau à une évaluation des différents besoins afin de planifier l'intervention (Macy, Thorndike-Christ et Lin, 2010). Malgré l'ensemble de ces critiques, les évaluations traditionnelles demeurent jusqu'à ce jour la méthode la plus utilisée afin de déterminer l'éligibilité des jeunes présentant potentiellement un RGD aux services d'intervention précoce (Macy *et al.*, 2015).

### 3. ÉLIGIBILITÉ AUX SERVICES D'INTERVENTION PRÉCOCES

Dans le présent mémoire, le terme éligibilité sera utilisé en référence aux critères sous lesquelles un enfant peut avoir accès aux services spécialisés d'intervention précoce. Bien qu'ils soient souvent associés, les critères d'éligibilité ne sont pas nécessairement les mêmes que ceux servant au diagnostic. En fonction des politiques en place, différents modèles peuvent être utilisés pour juger de la nécessité des services pour un jeune. Cette section présentera donc les critères utilisés aux États-Unis et au Québec.

#### 3.1 L'accès aux services spécialisés aux États-Unis et au Québec

Bien que la Individuals with Disabilities Education Act (IDEA; U.S. Department of Education, 2004) reconnaisse le retard de développement (RD) comme faisant partie des conditions nécessitant des services spécialisés d'intervention, le fait que les États-Unis demandent à chacun des 50 états de déterminer leurs propres critères d'éligibilité rend complexe la tâche de dresser un portrait exhaustif des politiques présentes dans ce pays. Le patron de critères permettant

l'éligibilité qui revient le plus souvent impose un retard de 25 % ou deux écarts-types dans un domaine de développement ou de 20 % ou 1,5 écart-type dans deux domaines de développement ou plus (Shackelford, 2006). Certains états fournissent également des services pour les enfants ayant un risque significatif de présenter un retard développemental (Shackelford, 2006). De plus, puisque les méthodes d'évaluation traditionnelles peuvent présenter plusieurs limites, quelques états ont décidé d'inclure la possibilité qu'une évaluation qualitative puisse mener aux services (Shackelford, 2006). Concrètement, cela signifie que le jugement clinique d'un professionnel peut être suffisant pour que l'enfant soit jugé éligible aux services. Cela ne concerne toutefois que quelques états, et ce, malgré le fait que le *Division For Early Childhood* (DEC, 2009) ait statué il y a déjà 10 ans sur l'importance d'utiliser des instruments rendant justice au développement de l'enfant dans son milieu naturel en plus d'affirmer qu'en aucun cas l'utilisation d'une seule mesure n'est souhaitable pour déterminer l'éligibilité d'un enfant.

Au Québec, chaque région possède un Centre Intégré de Santé et Service Sociaux (CISSS), ou un Centre Intégré Universitaire de Santé et Service Sociaux (CIUSSS), responsable d'offrir des services spécialisés de manière uniforme à tous les enfants présentant un RGD selon la définition du MSSS (2017) mentionnée ci-haut. Il est possible de remarquer que, bien que la définition du MSSS permette d'inclure un maximum d'enfants, elle ne se base que sur l'utilisation d'outils standardisés et n'implique pas le jugement clinique des différents professionnels. De ce fait, il est possible de mentionner que l'utilisation d'outils d'évaluation fonctionnels ou authentiques ne se fait pas de manière systématique pour chacun des enfants. Pour cette raison, il importe de développer des évaluations authentiques et valides pouvant être utilisées à cet effet.



### **3.2 Les services spécialisés d'intervention précoce québécois**

Tel que mentionné, ce sont les CI(U)SSS qui ont le mandat d'offrir des services spécialisés aux enfants présentant divers handicaps. Ces services exigent, de par la complexité des problématiques rencontrées, une grande expertise en matière du développement de l'enfant. Ils sont destinés aux enfants présentant les plus grandes difficultés sur le plan développemental et sont généralement dispensés par une équipe multidisciplinaire (psychoéducateur, ergothérapeute, orthophoniste, etc.; MSSS, 2017). L'objectif est de minimiser, dans un court délai, l'écart développemental entre un enfant présentant un RGD et les enfants au développement typique. En contrepartie, un enfant pour lequel il est jugé que des services spécialisés ne sont pas nécessaires peut tout de même recevoir des services plus généraux. Néanmoins, il est primordial d'identifier rapidement et correctement les enfants éligibles à ces services puisqu'une intervention précoce permet d'atténuer les conséquences du RGD et ce, autant à court terme qu'à long terme (Jonhson, 2016; Scherzer, Chhagan, Kauchali et Susser, 2012).

## **4. ÉVALUATION AUTHENTIQUE**

En réponse aux critiques véhiculées à l'endroit des évaluations traditionnelles et dans l'optique de favoriser l'évaluation rigoureuse des jeunes enfants présentant des difficultés développementales, la communauté scientifique en petite enfance s'est mise à la recherche d'une solution. L'utilisation d'évaluations authentiques axées sur l'observation et les activités quotidiennes et dont la passation s'effectue dans les milieux naturels des enfants est apparue

comme une alternative aux différentes limites exposées dans la section précédente (Bagnato *et al.*, 2010; Macy et Hoyt-Gonzales, 2007).

#### **4.1 Définition de l'évaluation authentique**

Il est possible de constater que plusieurs aspects reviennent dans la littérature lorsqu'il est question d'ÉA en intervention précoce. Les cinq critères qui suivent sont suggérés afin de déterminer ce qu'est l'ÉA. Le premier critère retrouvé dans la littérature est que les comportements évalués doivent être produits dans l'environnement naturel de l'enfant ainsi que dans ses routines quotidiennes (Bagnato et Ho, 2006; Dennis, Rueter et Simpson, 2013; Lazaro, 2017; Neisworth et Bagnato, 2004). Le second critère impose que l'évaluation soit faite auprès de ou par une personne qui est connue, significative et familière dans la vie de l'enfant ainsi que dans des milieux qu'il fréquente régulièrement (Macy, Bagnato et Gallen, 2016; Meisels, Wen et Beachy-Quick, 2010; Moreno et Klute, 2011). Troisièmement, lors de l'évaluation, l'enfant doit avoir suffisamment d'opportunités pour pouvoir démontrer de façon claire ses compétences et ses connaissances dans plusieurs domaines de développement et dans un contexte réel (Dennis *et al.*, 2013; Keilty *et al.*, 2009; Notary-Syverson et Losardo, 2004). Quatrièmement, différentes méthodes de collectes de données doivent être utilisées et différents acteurs significatifs de la vie de l'enfant doivent être consultés (Gao et Grisham-Brown, 2011; Neisworth et Bagnato, 2004; Notari-Syverson et Losardo, 2004). Le jugement clinique du professionnel entre également dans cette catégorie. Finalement, les résultats de cette évaluation doivent pouvoir servir à planifier des interventions ultérieures ou bien à déterminer les objectifs qu'il pourrait être pertinent d'inscrire au plan d'intervention. Ces données peuvent également servir à mesurer l'évolution d'un enfant en cours

d'intervention (Dennis *et al.*, 2013; Gao et Grisham-Brown, 2011; Keilty *et al.*, 2009; Macy et Bagnato, 2010). Dans un manuel expliquant ce qu'implique l'évaluation visant à planifier l'intervention, Bagnato *et al.* (2010) placent l'évaluation basée sur des curriculums comme figure première de l'évaluation visant l'intervention.

## 4.2 Les évaluations critériées du développement

Lorsqu'il est question d'instrument d'ÉA, il est souvent question d'instruments critériés et non d'instruments normalisés. Les évaluations dites *normatives* consistent à comparer les résultats d'un individu à la moyenne afin de le situer sur un continuum allant d'inférieur à supérieur à la moyenne. Pour leur part, les évaluations critériées utilisent un score seuil déterminé par les auteurs de l'instrument : si le développement du sujet se retrouve sous ce seuil, on juge que le trait mesuré est considéré problématique. Ainsi, l'utilisation de ce type d'évaluation est préférable lorsqu'il s'agit d'évaluer le développement d'un enfant présentant une problématique quelconque puisqu'elle permet de le comparer à une norme de fonctionnement établie plutôt que de le comparer à la moyenne des enfants ne présentant pas nécessairement de difficultés (Hallam *et al.*, 2004). À titre d'exemple, ces évaluations vont utiliser des activités de la vie quotidienne comme marcher, se laver les mains ou s'habiller comme items. Ce type d'instrument permettrait aux intervenants d'utiliser leur temps de manière plus efficace, d'impliquer les familles dans le processus, d'adapter l'instrument selon les besoins de l'enfant et d'identifier avec précision les besoins des enfants (de Sam Lazaro, 2017). En d'autres mots, il semblerait que les instruments critériés fournissent un portrait plus juste des enfants présentant des difficultés développementales (Squires, 2015).

### 4.3 Évaluation basée sur des curriculums

L'évaluation basée sur des curriculums (*Curriculum-based assessment*, ÉBC) se veut une forme d'ÉA misant principalement sur l'observation. Ce type d'évaluation met l'accent sur les comportements qu'il est nécessaire qu'un enfant soit en mesure d'émettre afin d'être indépendant dans son environnement (Macy *et al.*, 2010). La définition utilisée concernant l'ÉBC sera celle de Bagnato *et al.* (2010) et réfère à un cadre organisé dans lequel on retrouve ce que l'enfant doit apprendre, le processus par lequel il réalisera les objectifs identifiés et le changement dans ses performances à travers le temps. Un curriculum devrait contenir des informations détaillées sur 1) quoi apprendre, 2) comment l'enseigner, 3) où l'enseigner et 4) des moyens pour savoir si les enseignements sont fructueux.

Dans l'optique de décrire plus clairement en quoi consiste cette pratique, Notari-Syverson et Losardo (2004) stipulent que l'ÉBC comprend une série d'items ou de tâches servant à évaluer l'évolution d'un enfant en fonction des différentes étapes du développement. L'importance de recueillir des informations sur une longue période provient du fait que les jeunes enfants peuvent évoluer très rapidement sur une période de quelques mois. Cette rapide évolution peut avoir un impact significatif sur le développement d'un jeune ainsi que sur ses capacités adaptatives et donc influencer sur le statut d'éligibilité d'un jeune aux services spécialisés (Dhuey et Lipscom, 2010). Les résultats d'une ÉBC vont au-delà d'un score final, ils permettent de déterminer ce qu'un enfant est en mesure de faire ou de ne pas faire tout en prenant en considération les contextes de la vie quotidienne influant sur ses capacités (Bagnato *et al.*, 2010). Dans le contexte actuel, l'utilisation

des ÉBC pourrait permettre aux intervenants de procéder à la passation d'un outil leur permettant de corroborer une évaluation développementale traditionnelle tout en préparant l'intervention qui suivra pour un enfant.

#### **4.4 Questions de recension**

À la lumière de ce qui a été préalablement abordé dans la section de la problématique, les outils d'ÉA, et plus particulièrement les ÉBC, pourraient combler les différentes lacunes des évaluations traditionnelles du développement de l'enfant et ainsi améliorer le processus d'éligibilité des services d'intervention précoce. Toutefois, peu d'instruments permettent de corroborer le statut d'éligibilité d'un enfant, voire aucun disponible en français, et les paramètres entourant leur utilisation dans les milieux cliniques demeurent méconnus. De plus, la non-utilisation d'ÉA à des fins d'éligibilité s'explique généralement par l'absence de points de comparaison entre le sujet et la population générale. Ainsi, il serait pertinent d'examiner les outils d'ÉA ne présentant pas cette limite. Afin de dresser un portrait le plus juste possible des connaissances scientifiques sur l'évaluation du RGD de l'enfant, une recension des écrits a été réalisée pour tenter de répondre aux trois objectifs suivants :

1. Répertorier dans la littérature les évaluations authentiques permettant de corroborer l'éligibilité des enfants de 0 à 6 ans ayant un retard global de développement à recevoir des services spécialisés?

2. Documenter les caractéristiques psychométriques des évaluations authentiques permettant de corroborer l'éligibilité aux services d'enfants âgés entre 0 et 6 ans.
3. Documenter la nécessité que les intervenants utilisent des instruments d'évaluations authentiques permettant de corroborer l'éligibilité aux services d'enfants âgés entre 0 et 6 ans.

## DEUXIÈME CHAPITRE. LA RECENSION DES ÉCRITS

Cette section présente la recension des écrits réalisés. La méthode de recension ainsi que les résultats obtenus seront abordées. Finalement, l'objectif du mémoire sera précisé.

### 1. MÉTHODE DE RECENSION

Pour réaliser cette recension des écrits, cinq bases de données ont été consultées, soit PsycInfo, ERIC, MEDline, Academic Search et Education Source. Ces bases de données permettent de couvrir divers domaines de recherche (psychologie, éducation, médecine) pouvant être liés aux questions de recherches énumérées précédemment. Afin de bien cerner les études pertinentes, différents mots clés ont été sélectionnés à l'aide de la fonction *Thesaurus* des différentes bases de données. Trois concepts étaient visés par cette recension et des groupes de mots clés ont été identifiés pour chaque concept. Pour le premier concept, l'éligibilité aux services, les mots clés *Eligibility or "services accessibility"* ont été utilisés. Le second concept, les jeunes enfants présentant un RGD, a été sondé avec les mots suivants : (*"global developmental delay" or Developmental NI (delay or deviation) and "early childhood" or "child\*" or Early\**). Finalement, pour le troisième concept, les évaluations authentiques, deux groupes de mots clés liés par un *or* ont été utilisés. Le premier était : *"curriculum based assessment" or "authentic assessment"* alors que le deuxième était : (*"curriculum based assessment" or "authentic assessment"*) and (*"cutoff" or "AEPS" or "assessment, evaluation and programming system"*). Au total, quatre groupes de mots ont donc été utilisés pour les trois concepts. Chacune des bases de données a été sondée

individuellement et le seul critère mentionné était que les résultats de recherche devaient avoir été révisés par les pairs.

Le résultat de cette recherche a fait ressortir 87 articles qui se divisent comme suit parmi les bases de données : 18 pour PsycInfo, 0 pour MEDline, 20 pour ERIC, 16 pour Academic Search et 33 pour Éducation source. D'entrée de jeux, 46 articles ont été identifiés comme étant des doublons et retirés, portant ainsi le nombre d'articles à 41. Différents critères d'inclusion et d'exclusion ont ensuite été utilisés pour sélectionner les études les plus pertinentes. Pour être incluses, les études devaient être empiriques et concerner les enfants âgés entre zéro et six ans. Elles devaient également traiter de l'éligibilité aux services d'intervention précoce effectuée avec un ou des outils d'évaluation du développement jugés authentiques. Pour leur part, les outils d'évaluation destinés au trouble du spectre de l'autisme (TSA), les outils d'évaluation diagnostiques, les outils d'évaluation normatifs ainsi que ceux traitant d'un seul domaine de développement ont été exclus. Suite à la lecture des résumés de chacune des études, neuf ont été exclues car l'âge des participants était supérieur à six ans. Ensuite, 13 ont été exclues puisque le type d'outil d'évaluation ne convenait pas, c'est-à-dire que les instruments utilisés n'étaient pas réellement des ÉA. Finalement, neuf articles ont été exclus puisqu'il ne s'agissait pas d'études empiriques. La lecture des résumés a donc mené à la sélection de six articles pour une analyse approfondie. Toujours en fonction des critères ayant été mentionnés précédemment, quatre études ont été exclues. Deux ne présentaient aucun score seuil en lien avec l'éligibilité aux services, une était une recension des écrits et une avait des participants âgés de plus de six ans.



Finalement, seulement deux articles ont été retenus pour la présente recension. Ces deux articles (Bricker *et al.*, 2008; Bricker, Yovanoff, Capt et Allen, 2003) ont des objectifs très semblables à ceux du présent projet et ont la même première auteure. Afin d'élargir les résultats de la recension, certains critères d'exclusion ont été allégés. En effet, les articles non empiriques traitant directement de l'utilisation d'instrument d'évaluation authentique dans l'optique de corroborer l'éligibilité d'enfants âgés de moins de six ans aux services d'intervention précoce ont été retenus. De plus, les articles empiriques ayant d'autres objectifs que la détermination de scores seuils, mais traitant directement de la possibilité d'utiliser l'évaluation authentique dans l'optique de corroborer l'éligibilité de jeunes enfants aux services ont également été retenus. Finalement, deux recensions des écrits auparavant retirées des résultats en fonction des critères d'exclusion ont été considérées, compte tenu de leur pertinence. Sept articles supplémentaires ont donc été retenus, élevant le total à neuf. Toutefois, l'accès à un article est toujours impossible dû à un retard du prêt entre bibliothèques. L'ajout de ces articles a été fait dans le but de préciser le cadre théorique du présent projet ainsi que pour s'assurer de sa pertinence scientifique et clinique. Les résultats des études empiriques seront présentés distinctement de ceux des études non empiriques.

## 2. CARACTÉRISTIQUES DES ÉTUDES RECENSÉES

### 2.1 Études empiriques

Les deux études quantitatives retenues sont des études de validation exploratoires. C'est-à-dire que l'objectif principal vise à vérifier, par le biais de scores seuils, la possibilité d'utiliser l'EIS dans l'optique de corroborer l'éligibilité des jeunes enfants aux services d'intervention

précoce. Ensuite, l'objectif secondaire de ces études est de documenter la validité, la sensibilité et la spécificité des scores seuils établis.

Le Tableau 1 présente les principales caractéristiques méthodologiques de ces études. Il est possible de remarquer que la méthode est semblable pour les deux études. Au niveau de leurs échantillons, les deux études ont été menées avec des enfants ayant un développement jugé typique (n'étant pas éligibles aux services) ainsi que d'autres étant éligibles aux services. Celle de Bricker *et al.* (2003) est composée de 56,4 % de filles alors que celle de 2008 ne présente pas cette statistique. Afin de pouvoir réaliser la passation du protocole EIS, les enfants devaient avoir entre zéro et six ans, expliquant ainsi pourquoi les échantillons de ces études se retrouvent dans cette étendue. Une légère différence est observable en ce qui a trait à la version de l'outil ayant été utilisée. Dans l'étude de Bricker *et al.* (2003), la première édition de l'EIS (Bricker, 1993) a été utilisée pour les deux groupes d'âge, soit entre zéro et trois ans et entre trois et six alors que la deuxième édition de l'EIS (Bricker, 2002) a été utilisée uniquement pour le deuxième groupe d'âge. Dans la deuxième étude (Bricker *et al.*, 2008), seule la deuxième version a été utilisée pour les deux groupes d'âge.

Tableau 1. Caractéristiques des études empiriques

Articles	Échantillons				Outils utilisés		Procé- dure	Analyses		Résultats
	Dev. typi que	N Éligi bles	Tot.	Âge	Nom, Auteurs Années	Domaines évalués		Type analyses	Modèle d'analyse	
Bricker et al., 2003	603	258	861	0 - 6 ans	AEPS, Bricker , 2002, 1993	Motricité fine Motricité globale Adaptatif Cognitif Comm. sociale Social	Procéder à la passation de l'outil et additionner les buts cotés avec deux pour obtenir un score global.	Item Response Theory	Rasch Dichotomous One- Parameter Logistic  Modèle en 8 étapes	Cuttoffs scores à 2 écarts- types pour chaque tranche de 6 mois.
Bricker et al., 2008	719	662	1 381	0 - 6 ans	AEPS, Bricker , 2002	Motricité fine Motricité globale Adaptatif Cognitif Comm. sociale Social	Procéder à la passation de l'outil et additionner les buts cotés avec un ou deux pour obtenir un score global.	Item Response Theory	Rasch One- Parameter logistic partial credit model for polytomous scoring	Cuttoffs scores à 2 écarts- types pour chaque tranche de 3 mois.

## 2.2 Autres études sur le sujet

Différents types d'études ont été répertoriés parmi les six articles supplémentaires analysés. D'abord, deux recensions des écrits (Lazaro, 2017; Macy et Bagnato, 2010) sont ressorties. La première est présentée en première partie d'une étude plus large et vise à recenser les pratiques d'évaluation les plus fréquemment utilisées et de les comparer avec ce qui est recommandé par les organismes législatifs ainsi qu'avec les pratiques basées sur des données probantes. Pour ce faire, l'auteure présente les caractéristiques des évaluations traditionnelles et authentiques en plus d'aborder les forces et les limites des deux approches. Finalement, ces deux sections sont mises en relation avec ce qui est considéré comme étant les meilleures pratiques en matière d'évaluation du développement dans la littérature. La deuxième recension consiste, pour sa part, à examiner l'adéquation de neuf différents instruments d'évaluation authentiques sur le plan de la psychométrie (fidélité, validité et utilité). Au total, ce sont 27 études psychométriques, pour un total de plus de 10 000 enfants, qui ont été analysées. Pour être retenus, les instruments devaient avoir été soutenus par des recherches publiées dans des revues avec comité de lecture. L'AEPS (Bricker, 2002), la version anglophone de l'EIS fait partie de cette recension. Il s'agit de l'instrument répertorié ayant été l'objet du plus grand nombre de recherches ( $n = 9$ ). Ces deux études permettront de mieux comprendre l'importance et la pertinence d'inclure les instruments d'évaluation authentique dans le processus menant à l'éligibilité, mais aussi de documenter la capacité de ces instruments à identifier correctement le statut des enfants. D'ailleurs, une étude (Macy *et al.*, 2005) traite des caractéristiques psychométriques de l'EIS et s'attarde particulièrement à la capacité des scores seuils de la version américaine de l'instrument à identifier les enfants éligibles aux services d'intervention. L'étude, comprenant un échantillon de 68 dyades

parents / enfants, présente la valeur prédictive de l'instrument via des mesures de sensibilité, de spécificité, d'accord de classification, de validité concurrente et de fidélité interjuges. Pour ce faire, des comparaisons ont été effectuées avec différents instruments d'évaluation traditionnels ayant déjà été validés auparavant telles que le Gesell (Knobloch, Stevens et Malone, 1987), le *Batelle Developmental Inventory* (BDI; Newborg *et al.*, 1984, 2004), le Vineland (Sparrow *et al.*, 1984) et l'ASQ-2 (Bricker et Squires, 1999).

En plus d'une recension des écrits, l'étude de Lazaro (2017) a également utilisé un échantillon composé de 34 dyades donneurs de soin / enfants et de 14 praticiens pour documenter le processus d'évaluation actuel des jeunes enfants âgés entre zéro et trois ans. Pour ce faire, l'auteure a demandé aux intervenants de modifier leur pratique habituelle et ainsi de procéder à la passation d'un outil d'évaluation authentique peu importe le résultat de l'évaluation traditionnelle. L'objectif était de déterminer, par le biais d'entrevues et de questionnaires, si l'utilisation d'un instrument d'évaluation authentique basée sur un curriculum, le *Hawaii Early Learning Profile* (HELP; Parks, 2006), pouvait bonifier les informations obtenues par un instrument traditionnel ou si ces dernières étaient suffisantes. Bien que les données probantes suggèrent l'utilisation des deux types d'instruments (OPPQ, 2013), les résultats de cette étude pourront préciser le rôle des ÉA dans le processus d'éligibilité des enfants aux services d'intervention. Une autre étude (Keilty *et al.*, 2009) a utilisé des données recueillies auprès des intervenants réalisant les évaluations. Pour ce projet, réalisé aux États-Unis, 73 intervenants provenant de différentes ressources offrant des services d'intervention ont été rencontrés lors de sept *focus groups*. L'objectif de cette étude était de mettre en lumière les conditions idéales à l'utilisation de l'ÉA en termes de développement

professionnel et de gestion organisationnelle. Les analyses de théorisation ancrée (Glaser et Strauss, 1967) effectuée à partir des verbatims de chacun des *focus groups* visaient à documenter l'utilisation actuelle de l'ÉA ainsi que les facilitateurs et les barrières à sa mise en application dans les milieux cliniques américains. Bien que les résultats reflètent la situation actuelle des services américains et que cela soit difficilement transposable aux services québécois, ces derniers permettront de connaître la position des professionnels utilisant les outils d'évaluation.

Un autre type d'article, soit une étude de cas, a été retenu à la suite de la modification des critères d'exclusion. L'étude de Macy et Hoyt-Gonzales (2007) présente le parcours emprunté par une jeune fille de quatre ans et sa famille afin d'avoir accès à des services d'intervention dus à un retard de développement. L'objectif de l'article est de proposer une définition de ce que sont les ÉBC et de décrire leur rôle au sein du processus d'éligibilité des enfants. Pour ce faire, les auteures présentent chacune des étapes d'un modèle d'évaluation des jeunes enfants, le *Linked System Delivery System* (Pretti-Frontzak, 2002). Elles expliquent ensuite à l'aide de la situation mise en évidence par la famille participant à l'étude, comment chacune des étapes s'actualise dans la réalité clinique. L'article explique également en quoi les pratiques d'évaluation traditionnelles sont inefficaces et comment l'utilisation d'ÉBC peut devenir l'alternative permettant de combler bon nombre de leurs lacunes. Les impacts de l'utilisation des ÉBC sur l'ensemble de l'intervention, et non pas seulement sur l'étape de l'évaluation, sont également abordés.

Finalement, le dernier article retenu (McLean, 2005) se concentre sur les résultats présentés dans l'article de Macy *et al.* (2005) concernant les propriétés psychométriques de l'AEPS (Bricker,

2002). L'article présente l'évolution des pratiques d'évaluation en intervention précoce depuis le milieu des années 80 jusqu'au milieu des années 2000. L'auteur explique comment, malgré des lacunes importantes, les évaluations traditionnelles sont toujours utilisées de manière quasi universelle par tous les intervenants oeuvrant auprès de cette clientèle. Une deuxième section présente les résultats obtenus par Macy *et al.* (2005) et explique ce que les ÉA pourraient apporter aux différents services américains. Finalement, l'auteur conclut en élaborant sur l'importance d'un changement de paradigme dans le domaine de l'évaluation du développement des enfants.

### 3. LES RÉSULTATS DES ÉTUDES

Avant toute chose, il importe de spécifier que les articles analysés lors de cette recension contenaient tous des résultats traitant des différentes lacunes des instruments traditionnels d'évaluation du développement de l'enfant. Toutefois, comme ce sujet est abordé en profondeur dans la section problématique, seuls les résultats en lien avec l'ÉA sont rapportés afin d'éviter la redondance des informations ainsi que pour alléger le texte.

#### 3.1 Objectif 1 : les instruments développés

Seulement deux articles concernant le développement d'instrument d'ÉA dans l'optique de corroborer l'éligibilité aux services ont été recensés. Dans les deux cas, l'outil utilisé était le *Assessment, Evaluation and Programming System* (AEPS; Bricker, 2002). Puisqu'il s'agit de l'instrument de mesure utilisé dans le présent projet, il sera abordé en profondeur dans la section méthodologie. Pour la première étude (Bricker *et al.*, 2003) les résultats ont montré qu'il était

possible de créer des scores seuils pour chaque tranche d'âges de six mois et que ces derniers étaient en mesure d'identifier efficacement les enfants devant avoir accès aux services. La sensibilité permet d'évaluer la capacité d'un instrument à identifier adéquatement les enfants ayant une problématique, dans ce cas-ci un RGD, alors que la spécificité réfère à sa capacité à identifier correctement les enfants n'ayant pas de problématique. La sensibilité moyenne calculée est de 89,4 % et oscille entre 77 % pour la tranche d'âge de 67 à 72 mois et 100 % pour la tranche entre 55 et 60 mois. Pour sa part la spécificité moyenne est légèrement plus basse, mais demeure acceptable, à 76,4 %, et varie entre 70 % pour les tranches de 7 à 12, de 43 à 48 et de 49 à 54 mois jusqu'à 89 % pour l'intervalle de 67 à 72 mois. Les auteurs y ont également rapporté des mesures d'accord entre le statut déterminé par les évaluations traditionnelles et le statut déterminé par les scores seuils de l'instrument. En moyenne, 80,4 % [73 % - 91 %] des enfants ont reçu le même statut d'éligibilité lors de la passation de l'AEPS (Bricker, 2002) que lorsqu'ils ont été évalués lors de leur entrée dans les services. Les auteurs expliquent que les scores de spécificité moins élevés pourraient refléter le fait que les enfants évalués par le biais d'instruments standardisés traditionnels sont souvent considérés comme non éligibles alors que leur fonctionnement est différent comparativement à leurs pairs ne présentant pas de RGD et que des services leur seraient bénéfiques. Il ne s'agirait donc pas d'un manque de précision des scores seuils calculés.

La seconde étude (Bricker *et al.*, 2008) utilise elle aussi l'AEPS et se veut une réplique, plus approfondie, de la première étude. Tel que mentionné précédemment, la principale différence entre les deux articles est l'utilisation d'intervalles d'âge de trois mois pour la deuxième étude, comparativement à six mois pour la première étude. Les intervalles de trois mois ont été utilisés



pour les enfants âgés entre 0 et 36 mois. Les intervalles pour les enfants âgés entre 37 et 66 mois demeurent identiques à la première étude, soit des intervalles de six mois. Plusieurs informations statistiques supplémentaires sont également présentées dans cette étude. Premièrement, des corrélations entre l'âge des participants et les scores obtenus à l'AEPS (Bricker, 2002) ont été effectuées. Celles-ci varient entre .85 et .90 (0 à 36 mois) ainsi qu'entre .30 et .60 (37 à 66 mois) et sont toutes significatives ( $p < .01$ ). Deux autres indices psychométriques ont été calculées, soit la sensibilité et la spécificité. La sensibilité varie entre 76 % (43 à 48 mois) et 100 % (31 à 33 mois) avec une moyenne de 89,4 % alors que la spécificité varie entre 36 % (0 à 3 mois) et 80 % (28 à 30 mois) avec une moyenne de 66,7 %. Il est intéressant de noter que si la sensibilité calculée est identique en tout point à celle de l'étude de 2003, la spécificité est pour sa part plus basse de 10 %. Il en va de même pour le pourcentage d'enfants ayant reçu un statut d'éligibilité identique à l'instrument et à leur entrée dans les services puisqu'elle a légèrement diminué ( $M = 76,8$  %). Les auteurs notent toutefois des résultats très élevées de sensibilité et de corrélation (entre âge et score) pour les tranches d'âge de trois mois comparativement aux tranches d'âge de six mois, montrant ainsi que la division des scores seuils en de plus petits intervalles d'âge augmente leur efficacité.

### **3.2 Objectif 2 : les caractéristiques psychométriques des instruments**

Bien qu'elles ne traitent pas directement du développement d'un instrument d'ÉA visant à corroborer l'éligibilité, d'autres études se sont penchées sur les caractéristiques psychométriques d'instruments existants et sur leur capacité à distinguer les jeunes pour qui des services sont nécessaires de ceux pour qui le développement semble typique. Macy *et al.* (2005) se sont

d'ailleurs intéressés plus en profondeur aux caractéristiques psychométriques des scores seuils de l'AEPS proposés par Bricker *et al.* (2003). Dans cette étude, trois mesures de psychométrie étaient présentées. Premièrement, les résultats montrent un accord de classification moyen de 94 % lorsque les statuts d'éligibilité obtenus à l'AEPS (Bricker, 2002) sont comparés avec ceux du Gesell (Knobloch *et al.*, 1987) et du BDI (Newborg *et al.*, 1984). Les calculs de sensibilité et de spécificité sont, respectivement, de 100 % et de 89 %. Deuxièmement, la validité convergente indique des corrélations positives et significatives de 0,52 et 0,65 entre les scores totaux de l'AEPS (Bricker, 2002) et ceux obtenus au Vineland (Sparrow *et al.*, 1984) et au BDI (Newborg *et al.*, 1984). Compte tenu du fait que le Gesell (Knobloch *et al.*, 1987) ne comprend aucun score total, des corrélations ont été effectuées entre les scores obtenus aux différentes échelles des deux instruments. Celles-ci se sont avérées positives et significatives pour la majorité. Finalement, l'accord interjuges a été évalué et montre, pour le score total, une corrélation de Pearson de  $r = 0,92$  et un kappa de Cohen de  $\kappa = 0,93$ .

Des résultats psychométriques d'autres instruments d'ÉA ont également été obtenus par le biais de la recension de Macy *et al.* (2010). La typologie des études concernant les neuf instruments répertoriés est répartie comme suit : 13 études sur la fidélité interjuges, 5 études sur la fidélité test-retest, 5 études sur l'évaluation de la sensibilité et / ou de la spécificité, 15 études sur la validité convergente, 3 études sur la validité prédictive et finalement 6 études sur la validité de construit. La raison pour laquelle la somme des études est supérieure à 27, qui est le nombre d'articles retenus dans cette recension, est que plusieurs types d'études pouvaient se retrouver dans un même article. Parmi ces articles, 25 ont utilisé un échantillon d'enfants présentant des besoins

particuliers tel qu'un retard de développement ( $n = 10$ ), des problèmes de langage ( $n=4$ ) et des problématiques neurologiques ou génétiques ( $n = 7$ ) alors que cinq autres articles ont utilisé des enfants ayant été jugés éligibles aux services. Bien que l'AEPS soit l'outil regroupant le plus grand nombre d'écrits ( $n = 9$ ) traitant de ses caractéristiques psychométriques, plusieurs instruments sont concernés par au moins deux études, c'est notamment le cas pour l'*Adaptive Behavior Assessment System* (Harrison et Oakland, 2003;  $n = 2$ ), le *Child Observation Record* (High/Scope Educational Research Foundation, 2003;  $n = 3$ ), le *Developmental Observation Checklist System* (Hresko, Miguel, Sherbenou et Burton, 1994;  $n = 5$ ), le *Hawaii Early Learning Profile* (Parks, Furono, O'Reilly, Inatsuka, Hoska, et Zeisloft-Falbey, 1994;  $n = 2$ ), le *Pediatric Evaluation of Disability Inventory* (Haley, Coster, Ludlow, Haltiwanger et Andrellos, 1992;  $n = 2$ ) et finalement le *Transdisciplinary Play-Based Assessment* (Linder, 1993;  $n = 4$ ). En faisant référence aux nombreux résultats disponibles pour l'ensemble des instruments d'ÉA, les auteurs rapportent que la validité convergente, qu'ils expliquent être la meilleure façon de juger de la capacité d'un instrument à déterminer l'éligibilité d'un enfant aux services, a été calculée et jugée satisfaisante pour la majorité des instruments.

### **3.3 Objectif 3 : Pertinence des évaluations authentiques**

En plus de proposer une piste de solution visant à combler les lacunes des instruments traditionnels d'évaluation du développement de l'enfant, les ÉA ont de multiples impacts sur le processus d'évaluation des enfants. Parmi ceux-ci, la recension met en lumière l'augmentation du rapport temps-efficacité pour les intervenants (Bricker *et al.*, 2003; Lazaro, 2017) ainsi qu'une meilleure implication des milieux de garde et des familles dans l'évaluation (Macy et Hoyt-

Gonzales, 2007). Le fait que les ÉA soient développées pour permettre une évaluation de plusieurs domaines de développement simultanément pourrait expliquer cela (Bricker *et al.*, 2003; Macy et Hoyt-Gonzales, 2007). Ces résultats sont compatibles avec ceux de Keilty *et al.* (2009) qui indiquent que les intervenants responsables de l'évaluation des enfants reconnaissent l'importance d'utiliser des méthodes d'ÉA, notamment pour favoriser l'implication des familles et leur permettre une meilleure compréhension de l'intervention. De plus, certains intervenants mentionnent utiliser les stratégies d'ÉA pour procéder à la cotation de certains items provenant d'instruments traditionnels lorsque ceux-ci ne permettent pas une adaptation qui serait nécessaire à la condition de l'enfant (Keilty *et al.*, 2009; McLean, 2005).

Bien que le concept et l'expression *évaluation authentique* semblent inconnus de quelques intervenants, ces derniers affirment utiliser régulièrement cette pratique lors de l'ensemble des évaluations, qu'elles servent à déterminer l'éligibilité, à planifier l'intervention ou bien à évaluer les progrès d'un enfant. (Keilty *et al.*, 2009; Macy et Bagnato, 2010; Macy et Hoyt-Gonzales, 2007). D'ailleurs, les entrevues avec la famille et l'observation s'avèrent être les deux méthodes d'ÉA privilégiées par les intervenants lors du processus d'éligibilité (Keilty *et al.*, 2009). Les ÉA permettraient également aux cliniciens de porter un meilleur jugement clinique lorsque les résultats des évaluations traditionnelles du développement placent un enfant dans une zone à risque pour son développement ou bien lorsque celui-ci présente une condition déjà connue pouvant influencer sa performance (Bricker *et al.*, 2003). Cela est également vrai lorsqu'il est question des enfants ayant des différences socioculturelles telles que la langue ou la religion (Lazaro, 2017).

Même si les écrits scientifiques et législatifs mettent l'emphasis sur l'importance d'utiliser des outils d'ÉA tel que les ÉBC, les pratiques actuelles en matière d'évaluation du développement de l'enfant font état d'une toute autre réalité, notamment dû au fait que les instruments traditionnels sont bien implantés dans le système nord-américain (Bricker *et al.*, 2003; Lazaro, 2017; Macy et Hoyt-Gonzales, 2007). Ainsi, l'utilisation constante des ÉA dans les différents milieux cliniques permettrait d'offrir une évaluation basée sur des données probantes puisque, tel que mentionné dans le rapport du *Presidents' Commission on Excellence in Special Education* (Office of Special Education and Rehabilitative Services, 2002), l'utilisation courante des instruments traditionnels du développement de l'enfant devrait faire place à davantage de temps investi dans la cueillette d'informations fonctionnelles visant à planifier l'intervention (Bricker *et al.*, 2003).

Les cliniciens et les chercheurs reconnaissent la pertinence de l'ÉA pour combler le manque d'information rendue disponible pour l'intervention à la suite de la passation d'un instrument traditionnel (Bricker *et al.*, 2003; Keilty *et al.*, 2009). Ils ajoutent qu'une variété d'instruments d'ÉA est utilisée pour, entre autres, prioriser les objectifs, les opérationnaliser ainsi que pour déterminer les stratégies d'interventions pertinentes. L'utilisation préalable d'ÉA, particulièrement d'ÉBC, constitue une importante valeur ajoutée lors du développement des plans d'intervention puisque les résultats des évaluations peuvent être directement réutilisés comme moyens ou comme objectifs (Bricker *et al.*, 2003; Keilty *et al.*, 2009; Macy et Hoyt-Gonzales, 2007; McLean, 2005). Ainsi, le rendement des différents programmes d'intervention et l'utilisation judicieuse des ressources disponibles pourraient se voir augmenter par la mise en place d'une approche d'évaluation permettant de planifier efficacement l'intervention (Macy et Bagnato,

2010; Macy et Hoyt-Gonzales, 2007). De plus, bien que l'implication accrue des parents et des familles dans le processus d'évaluation ait été nommée plus haut, il en va de même au niveau de l'intervention. L'utilisation d'ÉBC permettrait aux parents de s'engager plus facilement dans l'intervention, notamment parce que les résultats touchent des habiletés fonctionnelles et observables dans la réalité quotidienne des familles (Macy et Hoyt-Gonzales, 2007). Finalement, les intervenants affirment également que les ÉA permettent d'effectuer des recommandations cliniques aux familles dont les enfants ont été jugés non éligibles à recevoir des services (Keilty et al., 2009).

### **3.4 Les recherches futures proposées**

Parmi les barrières identifiées par les intervenants, certains mentionnent que l'absence de mesures standardisées associées aux ÉA peut constituer un frein à leur utilisation puisque, d'un point de vue législatif, ces mesures sont nécessaires pour qu'un enfant puisse bénéficier de services (Keilty *et al.*, 2009). Ainsi, plusieurs études soulèvent l'importance de continuer la recherche sur l'utilisation de l'ÉA en s'attardant plus particulièrement à la précision de ces outils (validité et fidélité) ainsi qu'en les comparant à des standards déjà établis (Lazaro, 2017; Macy et Bagnato, 2010; McLean, 2005). En ce sens, Bricker *et al.* (2003) affirment que le développement futur d'instruments d'ÉA permettant de déterminer ou de corroborer l'éligibilité aux services de manière efficace s'avèrerait un progrès majeur dans le domaine de la petite enfance. Concernant la méthodologie, McLean (2005) mentionne qu'il serait intéressant d'inclure des enfants présentant de très faibles, mais significatifs, retards développementaux afin de déterminer si l'EIS est en mesure de détecter adéquatement ces enfants. Il mentionne également qu'il serait préférable que

les scores seuils soient déterminés pour chacun des domaines de l'outil, ce qui a été effectué lors de l'étude de Bricker *et al.* (2008).

### 3.5 Synthèses des études recensées

Afin de bien comprendre les résultats de la recension, les trois objectifs de celle-ci seront d'abord rappelés. Cette recension visait à répertorier les instruments d'ÉA permettant de corroborer l'éligibilité aux services spécialisés des jeunes enfants, à documenter les caractéristiques psychométriques des ÉA et finalement à statuer sur la pertinence d'utiliser des ÉA dans l'optique de corroborer l'accès aux services spécialisés des jeunes enfants.

Les résultats identifiés concernant les outils d'ÉA pouvant être utilisés à des fins d'éligibilité se sont avérés très homogènes puisqu'un seul instrument a été recensé. Bien que ce constat rende impossible la comparaison des outils entre eux, cela apporte tout de même son lot d'informations pertinentes puisque l'instrument utilisé est la version anglophone de l'EIS (Bricker, 2006) qui servira au présent projet. En effet, les statistiques plus qu'acceptables présentées au niveau de la sensibilité, de la spécificité et du degré d'accord de classification permettent de croire que la méthode utilisée par les auteurs s'est avérée efficace et ainsi, qu'elle peut servir de modèle à une autre étude similaire. La différence dans les résultats à la suite de la segmentation des tranches d'âge dans la deuxième étude permet également de conclure que de petits groupes d'âge devraient être utilisés. Bien que les caractéristiques psychométriques de l'AEPS (Bricker, 2002) soient documentées dans de nombreuses études, il aurait été intéressant de se pencher sur la fidélité interjuges de ces cueillettes de données. Même si cette statistique n'est pas présentée, les auteurs de l'étude de Bricker *et al.* (2008) mentionnent, en revanche, que la passation de l'instrument a été

effectuée par des professionnels ayant reçu une formation appropriée et pour lesquels la fidélité interjuges a été établie. Afin de combler ce manque d'information, il est possible d'utiliser l'étude de Macy *et al.* (2005). Celle-ci permet de cerner avec plus de précision les qualités psychométriques associées aux scores seuils, mais aussi à l'instrument en général.

Il a ensuite été possible de constater que les impacts positifs sont nombreux lorsqu'il est question d'utiliser l'ÉA à des fins d'éligibilité, et ce, tant au niveau de l'évaluation que de l'intervention. Ces impacts touchent diverses sphères de la vie des enfants affectés par diverses problématiques, mais aussi leurs parents ainsi que leurs intervenants. Parmi les plus notables, on remarque une augmentation de l'efficacité des intervenants ainsi que l'amélioration de l'implication des familles dans l'intervention. Il s'agit d'un impact majeur puisqu'une bonne participation des parents dans l'intervention constitue un facteur clé permettant d'atteindre les différents objectifs fixés (Keilty *et al.*, 2009; Neisworth et Bagnato, 2014). L'impact le plus important demeure néanmoins la possibilité de voir le rendement et l'efficacité des services dispensés aux enfants présentant diverses incapacités être améliorés par l'utilisation des ÉA. Il doit aussi être souligné qu'aucun inconvénient important n'a été soulevé concernant l'utilisation des ÉA. Toutefois, peu d'études se sont intéressées à cibler les facteurs négatifs. Les intervenants ont reflété deux difficultés à son utilisation. Dans un premier temps, certains sentent qu'ils n'ont pas suffisamment de connaissances pour réaliser une évaluation visant à déterminer le niveau développemental d'un enfant dans l'ensemble de ses domaines de développement, les professionnels ayant des domaines précis selon leur formation semblent être particulièrement touchés (ex : orthophonistes). Dans un second temps, le fait que l'ÉA puisse obliger les



intervenants à effectuer des recommandations sur les pratiques éducatives des parents semble créer un malaise chez certains intervenants.

Bien que l'ÉA fasse l'objet de plus en plus de projets de recherche, bon nombre d'aspects liés à son utilisation doivent encore être analysés. D'abord, aucun article ne fait état d'études ayant été réalisées auprès de participants québécois ou même francophones et ce, non seulement pour les scores seuils, mais aussi pour les instruments de manière générale. De plus, bien qu'il soit possible de documenter les qualités psychométriques de plusieurs instruments d'ÉA en consultant la littérature, aucun article ou projet, à l'exception des deux études citées ci-haut, n'a tenté de déterminer des scores seuils visant à déterminer l'éligibilité des jeunes enfants. Les écrits sont toutefois clairs, la communauté scientifique doit continuer à mener des projets de recherche allant en ce sens dans l'optique d'offrir le plus d'options possibles permettant de combler les diverses lacunes des instruments traditionnels.

En somme, les trois objectifs de recension mentionnés ont été partiellement atteints, notamment en raison du nombre restreint d'études concernant directement le développement d'instruments. En effet, il aurait été intéressant d'avoir accès à des ouvrages scientifiques ayant tenté de déterminer des scores seuils pour d'autres outils que l'EIS et avec d'autres analyses statistiques. Bien que les questions traitant davantage de la méthodologie n'aient pas été répondues avec précision, la littérature disponible a permis de bien répondre à la question 3 concernant la pertinence clinique des scores seuils. Les résultats, tels que précédemment mentionnés, témoignent de la pertinence clinique du projet puisque les retombées sur les résultats de l'intervention auprès des jeunes enfants sont positives. De plus, la possibilité de minimiser les conséquences liées aux

limites des instruments traditionnels et, par le fait même, de rendre les services destinés aux jeunes enfants plus performants soutient la pertinence sociale du projet. Finalement, au niveau de la pertinence scientifique, il est possible d'affirmer que le développement d'instruments d'ÉA permettant de corroborer l'éligibilité aux services constitue un besoin cité dans la littérature. Il est espéré que le présent projet permettra de contribuer à la réflexion concernant les questions 1 et 2 de cette recension.

#### 4. RETOMBÉES POUR LA PSYCHOÉDUCATION

Afin d'encadrer les psychoéducateurs œuvrant en petite enfance, l'OPPQ a développé des lignes directrices concernant l'évaluation d'enfant présentant potentiellement un RGD. Celles-ci réitèrent que les évaluations fonctionnelles sont primordiales puisqu'elles permettent de saisir la signification des différents comportements de l'enfant et leurs impacts sur les capacités adaptatives (OPPQ, 2013). Elles ajoutent que, par son expertise et son identité professionnelle, le psychoéducateur est en mesure d'utiliser le vécu partagé et l'observation participante pour évaluer qualitativement le développement d'un enfant dans des délais rapides et d'utiliser des outils de dépistage afin d'identifier les enfants nécessitant une évaluation approfondie. De plus, il est possible de poser l'hypothèse que l'arrivée du RGD comme trouble au sein du DSM-5 aura comme effet que les psychoéducateurs seront davantage sollicités pour contribuer à l'évaluation diagnostique du trouble effectué par d'autres professionnels (pédopsychiatre, médecin, neuropsychologue, etc.). Cette pratique est de plus en plus observée dans les milieux de pratique

puisque le psychoéducateur est en mesure de mettre en évidence des liens entre les difficultés adaptatives vécues au quotidien par un enfant et les symptômes apparentés au RGD (OPPQ, 2015).

À la lumière de ces informations, il est permis de croire que l'accessibilité d'évaluations authentiques permettant de corroborer l'éligibilité de jeunes enfants aux services spécialisés permettrait d'optimiser l'intervention psychoéducative auprès de cette clientèle. Dans un premier temps, la fluidité des opérations professionnelles que constituent l'observation, l'évaluation pré-intervention, la planification et l'évaluation post-intervention (Gendreau, 2001) se verrait augmentée puisque l'évaluation authentique utilise l'observation des comportements fonctionnels de l'enfant non seulement pour évaluer la nécessité de services spécialisés, mais également pour proposer des cibles et des stratégies d'intervention dont les résultats à moyen terme peuvent être comparés à ceux de l'évaluation initiale. Dans un deuxième temps, les psychoéducateurs pourraient contribuer davantage aux évaluations réalisées en équipe multidisciplinaire puisqu'ils seraient en mesure de statuer, en s'appuyant sur des évaluations valides, sur la nécessité des services spécialisés en fonction des capacités fonctionnels d'un enfant en interaction avec son environnement. Finalement, les évaluations authentiques pourraient permettre aux psychoéducateurs de conserver la spécificité de leur identité professionnelle par la réalisation d'une évaluation rigoureuse des capacités adaptatives basées sur des données probantes tout en conservant une proximité avec l'intervention.

## 5. OBJECTIFS INITIAUX POURSUIVIS

Le projet de recherche vise à évaluer la capacité des scores seuils de la version américaine de l'EIS à déterminer si un enfant est éligible ou non aux services spécialisés d'intervention précoce québécois. Pour se faire, les résultats de 50 enfants aux domaines cognitif et motricité fine seront mis en relation avec les scores seuils. Les résultats qui seront obtenus permettront d'amorcer une piste de réflexion concernant l'utilisation des instruments d'ÉA dans l'optique de corroborer l'éligibilité d'enfants présentant un RGD aux services d'intervention précoce. Premièrement, il sera possible de documenter la pertinence d'utiliser un instrument d'évaluation authentique dans une optique d'accès aux services d'intervention précoce. Deuxièmement, les résultats contribueront au développement de la littérature concernant les propriétés psychométriques des ÉA en contexte québécois.

## 6. CONTEXTE PARTICULIER DU MÉMOIRE

Dans un premier lieu, différents problèmes de recrutement ont été rencontrés lors de la collecte de données. Afin de pallier ce manque de participants, au printemps 2019, le nombre de participants a été diminué à 50, soit 25 enfants en CIUSSS et 25 enfants en CPE. Malgré cette modification, le recrutement des enfants admissibles aux services, c'est-à-dire les enfants présentant un RGD ayant accès aux services spécialisés du CIUSSS, demeurait très difficile. Une modification subséquente a donc eu lieu à l'automne 2019. Celle-ci visait à modifier les méthodes de recrutement pour procéder plutôt à une analyse de dossiers rétrospective. Cela facilite le recrutement puisque le consentement des usagers n'est pas nécessaire lors de cette procédure. Toutefois, cela nécessite une modification au niveau de la certification éthique, tant au CIUSSS de la Mauricie-et-du-

Centre-du Québec qu'à l'Université de Sherbrooke. Les confirmations du comité éthique de l'Université de Sherbrooke ont été reçues à l'hiver 2020. Deux demandes de convenances institutionnelles ont ensuite été réalisées, une au CISSS de Lanaudière et l'autre au CIUSSS du Saguenay-Lac-Saint-Jean. Ces centres ont été choisis puisque des ententes de collaboration entre les infrastructures de recherche dans lesquels le responsable du projet est impliqué facilitent le recrutement dans ces milieux. Les autorisations de convenance ont été reçues au printemps 2020. Toutefois, les mesures sanitaires mises en place suite à la déclaration de pandémie liée à la COVID-19 ont rendu la consultation des dossiers impossible puisqu'ils doivent être consultés en établissement. L'incertitude quant au délai d'attente a donc mené aux modifications suivantes concernant le projet.

Compte tenu de l'accessibilité restreinte à une partie des données nécessaires à la réalisation du projet dans sa totalité, diverses pistes de solutions ont été envisagées. La solution retenue est de considérer uniquement les enfants non admissibles aux services et pour lesquels les données avaient déjà été collectées par l'auteur de ce mémoire (lors de la première collecte de données) et d'abandonner le deuxième groupe d'enfants. Cependant, ce choix a entraîné une diminution significative de l'échantillon, ce dernier passant de 50 enfants pressentis à 21 enfants évalués. Les objectifs devaient également être révisés. En effet, il est impossible de comparer les deux groupes d'enfants compte tenu du retrait d'un des deux groupes, soit celui qui était composé des enfants éligibles aux services. Il est toutefois possible de faire une comparaison entre les deux instruments de mesure utilisés avec le groupe disponible. Ainsi, le nouvel objectif de ce projet de mémoire s'intéresse à la distance entre les scores obtenus par chaque enfant et le score seuil. Cette nouvelle

façon de faire permet l'utilisation d'un seul groupe d'enfant, ainsi que la comparaison des scores obtenus aux instruments. Cette comparaison sera réalisée en utilisant la distance entre les scores obtenus par les enfants et les scores seuils des deux instruments. Bien que les instruments ne possèdent pas le même nombre d'items, cette méthode fait en sorte qu'il est possible de standardiser les résultats obtenus en divisant la distance par le nombre d'items réalisés par l'enfant. De cette façon, il est possible d'explorer si les deux instruments (EIS et ASQ-3) ont la capacité d'identifier adéquatement les enfants pour lesquels aucun retard de développement n'est apparent. L'objectif de recherche se formule ainsi :

- Comparer la distance entre le score obtenu par l'enfant et le score seuil pour les domaines motricité fine et résolution de problème à l'ASQ et la distance entre le score obtenu et le score seuil pour les domaines cognitif et de motricité fine de l'EIS.

## **TROISIÈME CHAPITRE. LA MÉTHODE**

### **1. CONTEXTE DE L'ÉTUDE**

La méthode du présent projet vise à identifier les différents paramètres qui seront utilisés afin de contribuer au développement des connaissances concernant les caractéristiques psychométriques de l'EIS. Les sections suivantes présenteront la méthode utilisée pour obtenir les résultats.

### **2. LE DEVIS DE L'ÉTUDE**

La présente étude utilise un devis méthodologique (Fortin, 2010), c'est-à-dire un devis s'intéressant aux caractéristiques psychométriques d'un instrument. La réalisation du projet comprend quatre étapes principales, soit 1) le recrutement, 2) la passation des instruments et la consultation de dossier, 3) les analyses et 4) la comparaison entre les instruments. La section Déroulement présentera en profondeur ces étapes.

### **3. PARTICIPANTS**

Au total, 21 enfants non éligibles aux services spécialisés (ayant un développement typique). Le nombre de participants a été déterminé en fonction du nombre minimal requis pour procéder à l'analyse des données initialement prévue, ce nombre étant de 25 enfants. Les parents de tous les enfants non éligibles aux services impliqués dans le projet ont signé un formulaire de consentement.

### 3.1 Sélection des enfants non éligibles aux services

Un enfant non éligible aux services correspond à un jeune ne recevant pas de services liés à un retard de développement et dont le score au *Age & Stage Questionnaires* (ASQ-3, Squires et Bricker, 2009) n'a identifié aucun retard significatif. Pour effectuer le recrutement de ces enfants, deux centres de la petite enfance (CPE) de la Mauricie ont été contactés par l'entremise du Regroupement des centres de la petite enfance de la Mauricie et du Centre-du-Québec. Le projet a ensuite été présenté à la direction de ces deux établissements. Ces enfants ont été recrutés sur une base volontaire à la suite d'une description du projet aux parents de la part de la direction du CPE. Les enfants jugés non éligibles mais recevant des services pour d'autres problématiques pouvant porter atteinte à leur courbe développementale (ex : TSA, trouble du langage, etc.) ont été exclus du projet.

## 4. INSTRUMENTS DE MESURE

Deux instruments de mesure et un questionnaire sont utilisés pour recueillir les informations nécessaires à l'évaluation des scores seuils. D'abord, un questionnaire sert à amasser des renseignements sociodémographiques. Ensuite, l'ASQ-3 (Squires et Bricker, 2009) sert à confirmer que les enfants non-éligibles ne présentent pas de particularités développementales. Finalement, les domaines « motricité fine » et « cognitif » de l'EIS-2 (Bricker, 2006) sont administrés aux participants non éligibles par le responsable du projet.



#### **4.1 Questionnaire sociodémographique**

Un court questionnaire de renseignements sociodémographiques est rempli par les parents des enfants non éligibles participants à la suite de la signature du formulaire de consentement. Ce questionnaire, créé par l'auteur du mémoire, sert à recueillir certaines données nominatives dans l'optique de mieux décrire l'échantillon de participants. Ainsi, on y retrouve des questions sur la date de naissance de l'enfant, le sexe de l'enfant, les diagnostics actuels et leur date d'obtention (le cas échéant), le revenu annuel des parents et la fratrie des enfants inclus dans l'étude.

#### **4.2 Le Age & Stages Questionnaires 3 (Squires et Bricker, 2009)**

L'ASQ-3 est un outil de dépistage des problèmes développementaux chez l'enfant. Il permet, entre autres, l'identification des enfants pour lesquels un risque de RGD est présent. Le questionnaire est complété par les parents ou un proche de l'enfant et il est disponible pour les jeunes âgés entre 1 et 66 mois. L'ASQ est divisé en cinq domaines : la motricité fine, la motricité globale, les aptitudes individuelles et sociales, la résolution de problèmes et finalement la communication. Plusieurs motifs ont guidé le choix de cet outil. Il s'agit d'un questionnaire court, ne comptant que six items par domaine se répondant par « oui », « pas encore » ou « parfois » et dont l'utilisation n'est pas réservée aux professionnels. Cela permet donc de le distribuer aux parents ou aux éducateurs sans que la réalisation de l'outil ne leur prenne trop de temps. De plus, les items ont été créés pour assurer la compréhension d'un répondant ayant un niveau de lecture comparable à la quatrième ou cinquième année du primaire et des images sont disponibles pour faciliter l'interprétation des items. La simplicité du questionnaire permet donc aux parents de le

remplir facilement à la maison et ce, même lorsqu'ils présentent certaines limitations légères. Les familles qui avaient des questions ou désiraient obtenir du soutien pour la complétion du questionnaire pouvaient en obtenir en contactant le responsable du projet par téléphone ou bien par courriel, les informations à cet effet figuraient sur le formulaire de consentement préalablement signé.

L'ASQ-3 est composé de 21 questionnaires permettant de couvrir les différentes tranches d'âge. Il contient des questionnaires pour chacun des intervalles de deux mois pour les enfants âgés entre zéro et 24 mois (13 questionnaires)<sup>1</sup>, pour chacun des intervalles de trois mois pour les enfants âgés entre 27 et 36 mois (4 questionnaires) et finalement, pour chacun des intervalles de six mois pour les enfants âgés entre 42 et 66 mois (4 questionnaires). Les différentes propriétés psychométriques soutiennent l'utilisation de l'ASQ-3. Dans un premier temps, l'outil tient compte des enfants nés prématurément en ajustant l'âge lors du choix du questionnaire. De plus, le tiers des questions (soit deux questions sur six) peuvent être manquantes sans que le questionnaire ne devienne impossible à interpréter. Les différentes procédures à cet effet sont présentées dans le manuel de l'utilisateur de l'instrument. Les propriétés psychométriques de l'ASQ-3 sont bien documentées. Les résultats psychométriques qui suivent sont tirés du manuel de l'utilisateur (Squires, Twombly, Bricker et Potter, 2009) et sont basés sur la complétion de 18 572 questionnaires auprès de 15 138 enfants américains. La fidélité test-retest de l'instrument est de 92 % avec des corrélations intraclass variant entre 0,75 et 0,82 selon les domaines, ce qui est

---

<sup>1</sup> Le total est de 13 puisque des questionnaires distincts ont été créés pour les enfants âgés de 9 et 10 mois.

considéré comme une bonne fidélité interjuges (Koo et Li, 2017). La fidélité interjuges a été évaluée en comparant les réponses des professionnels avec celles des parents. Les résultats ont montré un accord interjuges de 93 %. La cohérence interne montre pour sa part des corrélations entre 0,60 et 0,85 (toutes significatives à  $p < 0,01$ ). Des alphas de Cronbach entre 0,51 et 0,87 ont également été calculés, ce qui peut être considéré entre acceptable et bon selon les domaines (Taber, 2018). La validité concurrente de l'instrument a été évaluée en comparant des résultats avec ceux obtenus aux deux éditions du *Batelle Developmental Inventory* (BDI; Newborg, Stock, Wnek, Guidubaldi, et Svinicki, 1984, 2004). Les résultats montrent que les résultats des deux instruments sont identiques pour 86 % des participants. L'ensemble des données disponibles ainsi que les caractéristiques de l'outil précédemment mentionnées expliquent le choix de cet instrument pour confirmer le statut développemental des participants au projet de mémoire.

#### **4.3 Le Programme Évaluation, Intervention, Suivi (Bricker, 2006)**

L'EIS fait partie des évaluations authentiques (Macy *et al.*, 2005) et répond aux cinq critères énumérés auparavant concernant ce type d'évaluation. Plus précisément, l'EIS s'inscrit dans la sous-catégorie des évaluations basées sur un curriculum (Waddell, Pretti-Frontzcak, Johnson et Bricker, 2007) et se veut une version francophone traduite de l'*Assessment, Evaluation and Programming System* (AEPS; Bricker, 2006). L'EIS est un outil d'évaluation critérié basé sur l'observation du comportement de l'enfant lorsqu'il agit dans son milieu naturel. Le Tome I est constitué du guide d'utilisation et du test, il comprend 464 items (249 pour le 0-3 ans et 215 pour 3-6 ans). Les enfants visés par l'EIS sont aussi bien ceux ayant un développement typique, mais étant à risque de RGD que ceux présentant un RGD ou tout autre trouble ou incapacité. L'outil

permet de dresser un portrait global et organisé de l'ensemble du développement et des comportements de l'enfant (Neisworth et Bagnato, 2004). L'EIS rend également plus facile le travail des équipes multidisciplinaires ayant fait le choix de l'utiliser puisqu'il permet la mise en place d'objectifs fonctionnels généralisables et compréhensibles pour l'ensemble des professionnels impliqués (Macy *et al.*, 2009). Le programme EIS permet de documenter le développement d'un enfant selon six domaines : motricité fine, motricité globale, adaptatif, communication, social et cognitif. Chaque domaine se divise en différents niveaux. Le nombre de niveaux varie selon les domaines. Ces niveaux permettent de regrouper certains items représentant une même action ou un même comportement. L'ordre dans lequel ils apparaissent a été déterminé de manière hiérarchique en fonction de la séquence développementale typique d'un enfant. Le niveau « A » est donc le plus facile et les suivants sont de plus en plus difficiles. À l'intérieur de ces niveaux se trouvent des buts et des objectifs suivant le même patron de présentation.

Afin de favoriser le lien entre l'évaluation de l'enfant et la planification de l'intervention, le Tome I présente des exemples de buts et d'objectifs destinés à la rédaction de plan d'intervention ou de services. Ceux-ci sont directement liés aux items de l'évaluation de l'EIS, ce qui permet aux intervenants de se repérer facilement dans les curriculums et d'identifier des objectifs correspondants au niveau actuel de développement de l'enfant. Les tomes II (0 à 3 ans) et III (3 à 6 ans) proposent pour leur part des interventions visant à travailler les différentes habiletés de l'enfant en lien avec les différents domaines. Chaque intervention est également reliée aux items de l'EIS tout en demeurant le plus naturaliste possible. Une autre section de ces curriculums porte

de manière plus spécifique sur des moyens d'intervention visant à permettre à un enfant d'atteindre certains buts du programme.

Dans les Tomes II et III, chaque domaine est présenté en détail. Pour chacun des niveaux, les auteurs expliquent l'importance des habiletés visées et suggèrent des interventions visant à faciliter l'observation des différents objectifs. Des précisions concernant de possibles modifications ou adaptations faites à l'environnement y sont décrites afin que les enfants ayant des besoins particuliers ou n'étant pas en mesure d'émettre un comportement puissent également être évalués. De plus, ces adaptations permettent aux intervenants de bien comprendre comment utiliser les accessoires en lien avec les objectifs (p. ex. une cuillère, une peluche ou une ficelle). Par la suite, une section nommée « Activités dirigées » propose certaines activités planifiées étant moins naturalistes. Bien que les observations liées à l'utilisation de l'EIS doivent idéalement se faire dans un contexte le plus naturel possible, certains enfants peuvent avoir besoin d'être encadrés davantage, notamment dû à leur condition développementale (Bricker, 2006). Finalement, on y trouve une présentation des différents aspects à considérer. Cette section concerne les précautions à prendre lorsque vient le temps de procéder à l'évaluation d'un enfant.

#### *4.3.1 Caractéristiques psychométriques de l'EIS*

Bon nombre d'études ont été réalisées afin d'évaluer les propriétés psychométriques de l'EIS et de l'AEPS. Les propriétés psychométriques rapportées dans cette section se référeront aux articles publiés à partir de l'année 2002, moment auquel la deuxième version de l'instrument fût publiée.

La fidélité test-retest a été évaluée dans une étude (Bricker, Bailey et Slentz, 1990) et les résultats ont montré une corrélation de 0,96 entre les deux temps de passation. La fidélité interjuges a, quant à elle, été l'objet de trois études (Grisham-Brown, Hallam et Pretti-Frontczak, 2008; Macy, Bricker et Squires, 2005; Noh, 2005). Les mesures de Kappa de Cohen varient entre 0,73 et 0,92 alors que les corrélations oscillent entre 0,70 et 0,93. La cohérence interne, évaluée par Noh (2005) à l'aide du Kappa de Cohen, varie entre 0,83 et 0,99.

En ce qui a trait à la validité, plusieurs auteurs ont comparé les résultats obtenus à l'EIS à ceux obtenus avec d'autres outils d'évaluation du développement. De manière générale, les enfants obtiennent des scores similaires lorsqu'ils sont évalués avec l'EIS que lorsqu'ils le sont avec d'autres outils bien reconnus (Slentz, 2008). Les études de validité convergente ont été réalisées avec des outils tels que, entre autres, le *Vineland Adaptive Behavior Scales* (Macy et al., 2005; Sparrow et al., 1984), le *Bayley Scale of Infant Development* (Bayley, 1969; Bricker et al., 1990) et le *Battelle Developmental Inventory* (Macy et al., 2005; Newborg, Stock, Wnek, Guidubaldi et Svinicki, 1988). La validité concourante a pour sa part été évaluée en corrélant les scores obtenus à différents domaines de l'EIS et du *Battelle Developmental Inventory, Second Edition* (Newborg et al., 1984, 2004). Les résultats se sont avérés positifs et significatifs (Newborg, 2005). L'instrument est également en mesure de différencier les enfants au développement typique de ceux ayant un développement atypique (Bricker et al., 2008).

Bagnato *et al.* (2014) se sont particulièrement intéressés à la validité sociale des outils d'évaluation authentique. Ils définissent la validité sociale comme étant l'appréciation subjective de la procédure d'évaluation par les personnes impliquées. Dans une étude visant à comparer l'ÉA aux évaluations standards, ces mêmes auteurs ont demandé aux intervenants de juger les outils d'évaluation selon huit critères (désirabilité, authenticité, collaboration, pertinence, multifactorialité, sensibilité, universalité et utilité) servant à définir les caractéristiques d'une évaluation développementale appropriée s'arrimant avec les meilleures pratiques d'intervention. Les résultats montrent que les ÉA répondent bien à ces critères. De plus, les parents d'enfants présentant ou non des difficultés préfèrent l'ÉA aux évaluations conventionnelles. Ils s'y considèrent plus impliqués et mentionnent qu'elle reflète davantage la réalité de leurs enfants (Macy *et al.*, 2010).

## 5. DÉROULEMENT

Le recrutement des participants a été réalisé à l'automne 2019. Pour sélectionner les enfants non éligibles et éligibles, plusieurs CPE de la Mauricie ont été contactés afin de leur présenter le projet. Les directeurs intéressés à participer au projet ont ensuite été invités à communiquer avec les parents des enfants fréquentant leur institution afin d'obtenir leur accord pour participer au projet. Une liste des numéros de téléphone des parents désirant prendre part au projet de recherche est ensuite acheminée à l'auteur de ce mémoire, qui communique alors avec chacun d'entre eux pour leur expliquer le projet. Un formulaire de consentement, un questionnaire sociodémographique ainsi que l'ASQ-3 leur a ensuite été remis en main propre par l'auteur ou par

un éducateur du CPE. Un court document présentant un résumé du projet ainsi que les coordonnées de l'auteur a été également distribué à tous les parents. Ces derniers étaient invités à contacter l'auteur s'ils avaient des interrogations ou s'ils désiraient davantage d'informations. Différentes plages d'observations ont ensuite été déterminées avec la direction des deux CPE participants pour que l'auteur puisse effectuer l'observation des enfants ayant obtenu un score se trouvant dans la normalité statistique à l'ASQ-3 et procéder à la cotation des domaines « motricité fine » et « cognitif ».

## 6. MÉTHODE D'ANALYSE DE DONNÉES

Dans un premier temps, les informations recueillies ont été compilées dans une base de données à l'aide du logiciel *IBM SPSS Statistics for Macintosh, Version 27.0*. Des statistiques descriptives ont ensuite été générées pour présenter l'échantillon. Des moyennes et des écarts-types sont ressortis pour l'âge ainsi que pour les scores à l'ASQ-3 et à l'EIS. Dans un troisième temps, la distance entre les scores obtenus et les scores seuils ont été calculés, puis standardisés en fonction du nombre de questions. Pour ce faire, la distance obtenue a été divisée par le nombre d'items répondus. Cela devait être fait puisque les deux tomes de l'EIS (0 à 3 ans et 3 à 6 ans) ne comportent pas le même nombre de questions. Finalement, pour répondre à l'objectif de recherche de ce mémoire, une corrélation entre les distances standardisées a été faite pour les deux domaines évalués dans le projet.



## QUATRIÈME CHAPITRE. LES RÉSULTATS

Ce chapitre vise à présenter les différents résultats obtenus suite aux analyses. Une première partie des résultats consistent en des statistiques descriptives permettant de décrire l'échantillon d'enfants ayant participé au projet. Les deux corrélations réalisées seront ensuite présentées, soit celle entre les scores obtenus au domaine de la motricité fine à l'ASQ et ceux obtenus au même domaine pour l'EIS ainsi que celle entre les résultats obtenus au domaine résolution de problèmes de l'ASQ et au domaine cognitif de l'EIS.

### 1. STATISTIQUES DESCRIPTIVES DES PARTICIPANTS

Les statistiques descriptives présentées concernent les enfants et leur famille. Elles sont tirées du questionnaire sociodémographique que les parents devaient remplir lors de la signature du formulaire de consentement. Au total, 21 enfants fréquentant deux milieux de garde de la Mauricie ont participé au projet. Une famille n'a pas retourné le questionnaire sociodémographique, les données présentées concernent donc 20 familles. Les enfants étaient âgés entre 10 et 60 mois au moment de l'étude alors que l'âge moyen des enfants était de 37,33 mois ( $\text{ÉT} = 12,85$ ). La moitié des familles rapportent avoir un revenu familial brut se situant entre 100 000\$ et 200 000\$ par an alors que huit familles se trouvent entre 60 000\$ et 100 000\$ et deux familles entre 40 000\$ et 60 000\$. La langue parlée à la maison est le français pour 18 des 20 répondants. Les deux autres familles affirment utiliser un mélange de l'anglais et du français à la maison. Finalement, 18

familles sont nucléaires alors qu'une est recomposée et qu'une autre mentionne avoir différentes modalités de garde.

## 2. COMPARAISON ENTRE LES INSTRUMENTS

La corrélation de Pearson a été utilisée pour vérifier la relation entre les variables concernées par la distance par rapport au score seuil de la motricité fine à l'ASQ et la distance par rapport au score seuil de la motricité fine à l'EIS. L'analyse statistique montre une corrélation non significative négative de  $-0,14$ ;  $p = 0,549$ . La même analyse a été utilisée pour vérifier la relation entre les variables concernées par la distance par rapport au score seuil de résolution de problème de l'ASQ et la distance par rapport au score seuil cognitif de l'EIS. L'analyse montre une corrélation non significative négative de  $-0,20$ ;  $p = 0,388$ . Les figures 1 et 2 présentent les diagrammes de dispersion des deux corrélations.

Figure 1. Diagramme de dispersion pour la motricité fine

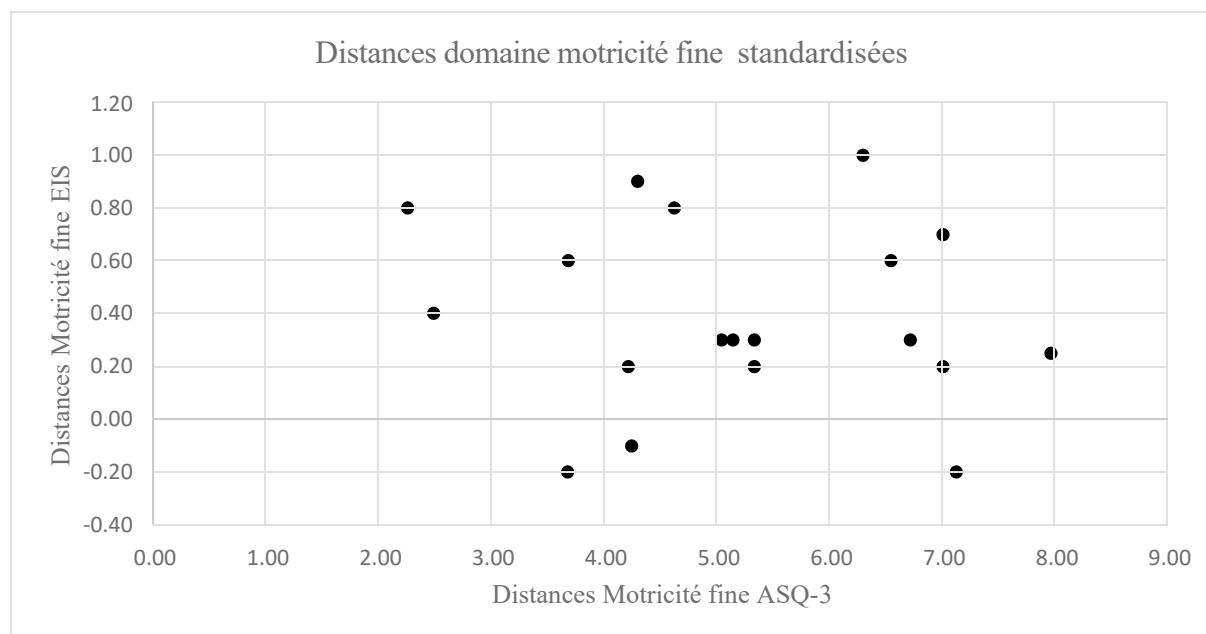
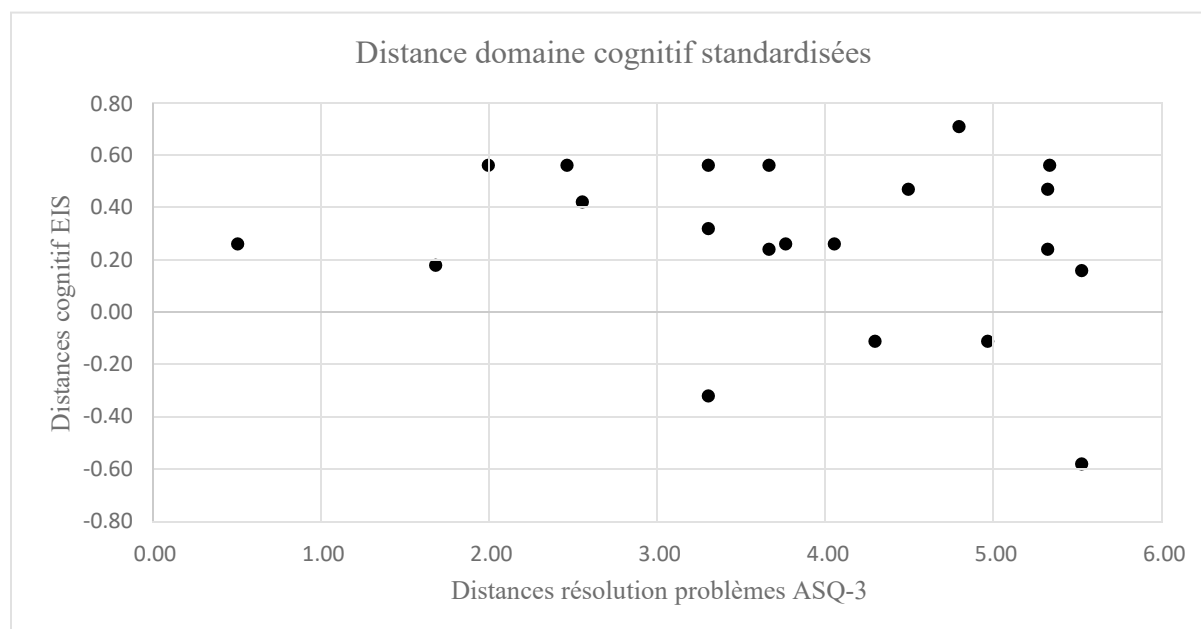


Figure 2. Diagramme de dispersion pour le domaine cognitif



Au total, sur les 42 résultats obtenus à l'EIS, sept résultats ne concordent pas avec les résultats obtenus à l'ASQ-3, c'est-à-dire que les deux instruments ne s'entendent pas sur la présence ou non d'un retard significatif pour un même domaine. Dans tous les cas, le résultat à l'EIS montre qu'un retard significatif est présent dans un des deux domaines alors que ce n'est pas le cas aux résultats de l'ASQ-3. Ces sept résultats touchent cinq enfants différents et six de ces résultats concernent le niveau 1 de l'EIS, soit la version zéro à trois ans. Ainsi, 83,3% (35 sur 42) des scores à l'EIS concordent avec les résultats obtenus à l'ASQ-3. L'ampleur des divergences entre les résultats indique que, pour le domaine de la motricité fine, les scores des enfants sont très près du score seuil. En effet, sur les trois résultats divergeant dans ce domaine, deux se trouvent à un point du score seuil alors que le troisième se trouve à deux points du score seuil. Pour les quatre résultats divergeant du domaine cognitif, les écarts sont plus grands. Deux enfants se trouvent à deux points du score seuil, alors qu'un autre se trouve à six points et le quatrième à onze points du

score seuil. Le tableau 1 présente les résultats obtenus pour chacun des enfants au domaine motricité fine de l'ASQ-3 et de l'EIS et le tableau 2 les résultats obtenus pour chacun des enfants au domaine cognitif. Il est possible de constater que quatre résultats divergents sur sept concernent des enfants âgés soit de 33 ou de 34 mois.

La distance entre les scores obtenus et les scores seuils varient également selon l'instrument et le domaine. On remarque que les scores obtenus à l'ASQ-3 se distancient davantage du score seuil que ceux de l'EIS et ce, même lorsque les scores sont standardisés en fonction du nombre d'items. En effet, les moyennes des distances à l'EIS sont de 0,27 pour le domaine cognitif et de 0,38 pour le domaine de la motricité fine. Les étendues sont de 1,29 et 1,20 respectivement. Au niveau de l'ASQ-3, les moyennes des distances calculées sont de 3,79 pour la résolution de problème et de 5,32 pour la motricité fine alors que les étendues sont de 5 et 5,71.

Tableau 2. Résultats au domaine motricité fine.

Participant	Âge en mois	Résultats ASQ-3	Score seuil ASQ-3	Distance standardisée ASQ-3	Résultats EIS	Score seuil EIS	Distance standardisée EIS
1	10.00	60	37.97	3.67	12	6.00	0.60
2	16.87	60	34.32	4.28	18	9.00	0.90
3	21.50	55	29.61	4.23	10	11.00	-0.10
4	23.23	50	35.16	2.47	16	12.00	0.40
5	28.45	50	19.25	5.13	16	13.00	0.30
6	32.80	60	12.28	7.95	18	15.50	0.25
7	33.00	60	12.28	7.95	18	15.50	0.25
8	34.30	40	18.07	3.66	14	16.00	-0.20
9	34.90	50	18.07	5.32	18	16.00	0.20
10	36.60	60	18.07	6.99	18	16.00	0.20
11	36.70	50	18.07	5.32	4	2.50	0.30
12	37.00	60	18.07	6.99	6	2.50	0.70
13	40.94	50	19.82	5.03	4	2.50	0.30

14	42.00	60	19.82	6.70	4	2.50	0.30
15	42.26	50	19.82	5.03	4	2.50	0.30
16	43.65	45	19.82	4.20	4	3.00	0.20
17	48.00	55	15.81	6.53	6	3.00	0.60
18	53.23	55	17.32	6.28	10	5.00	1.00
19	54.00	45	17.32	4.61	8	5.00	0.80
20	54.50	60	17.32	7.11	4	5.00	-0.20
21	60.10	40	26.54	2.24	10	6.00	0.80

Tableau 3. Résultats au domaine cognitif pour l'ASQ-3 et l'EIS

Participant	Âge en mois	Résultats ASQ-3	Score seuil ASQ-3	Distance standardisée ASQ-3	Résultats EIS	Score seuil EIS	Distance standardisée EIS
1	10.00	55	32.51	3.75	10	5.00	0.26
2	16.87	50	25.74	4.04	16	11.00	0.26
3	21.50	55	29.30	4.28	18	20.00	-0.11
4	23.23	45	29.78	2.54	28	20.00	0.42
5	28.45	30	27.08	0.49	30	25.00	0.26
6	32.80	55	26.92	5.51	32	29.00	0.16
7	33.00	60	26.92	5.51	18	29.00	-0.58
8	34.30	50	30.29	3.29	24	30.00	-0.32
9	34.90	60	30.29	4.95	28	30.00	-0.11
10	36.60	50	30.29	3.29	36	30.00	0.32
11	36.70	50	30.29	3.29	20	10.50	0.56
12	37.00	45	30.29	2.45	20	10.50	0.56
13	40.94	40	28.11	1.98	20	10.50	0.56
14	42.00	50	28.11	3.65	20	10.50	0.56
15	42.26	60	28.11	5.32	20	10.50	0.56
16	43.65	50	28.11	3.65	18	14.00	0.24
17	48.00	60	31.30	4.78	26	14.00	0.71
18	53.23	60	28.12	5.31	26	18.00	0.47
19	54.00	60	28.12	5.31	22	18.00	0.24
20	54.50	55	28.12	4.48	26	18.00	0.47
21	60.10	40	29.99	1.67	22	19.00	0.18

## **CINQUIÈME CHAPITRE. LA DISCUSSION**

Cette section du mémoire présente la discussion du processus de recherche et des résultats qui en découlent. L'objectif et le contexte du projet seront d'abord rappelés. Il s'en suivra une discussion sur l'interprétation des résultats obtenus. Finalement, l'implication des résultats pour la recherche sur l'évaluation du développement des jeunes enfants sera abordée de même que les limites de l'étude et les recommandations pour de futurs projets.

### **1. RAPPELDEL'OBJECTIFDUPROJET**

Le projet mené s'intéresse aux instruments d'évaluation du développement de l'enfant et plus particulièrement, aux instruments d'évaluation authentiques. Il découle d'un besoin énoncé dans la littérature concernant le développement d'instruments francophones valides. Ce projet s'inscrit également dans une tentative de pallier les différentes lacunes identifiées précédemment en lien avec l'utilisation d'instruments d'évaluation du développement standardisés lorsqu'utilisés à des fins de planification et de soutien à l'intervention. Pour ce faire, l'objectif du projet est de comparer la distance entre le score obtenu et le score seuil pour les domaines « motricité fine » et « résolution de problème » à l'ASQ et la distance entre le score obtenu et le score seuil pour les domaines cognitif et motricité fine de l'EIS.

### **2. COMPARAISON ENTRE LES DEUX INSTRUMENTS**

La comparaison entre les distances par rapport aux scores seuil de l'ASQ-3 et de l'EIS trouve sa pertinence lorsqu'on analyse la fonction des deux instruments. En effet, tel que

mentionné dans la section méthode, l'EIS est une évaluation basée sur un curriculum. Ce type d'évaluation est généralement utilisé dans le but de planifier l'intervention et demande du temps ainsi qu'une bonne connaissance de l'enfant (Bagnato, 2010). De son côté, l'ASQ-3 est un instrument de dépistage ayant une procédure de passation simple et accessible à tous, en plus de permettre une complétion rapide (Tardif, Bérubé, Coutu, Dubeau, Jacques, et Lemay, 2015). C'est donc dire que l'ASQ-3 demande moins de temps et de ressources à compléter que l'EIS. Ainsi, il pourrait être avantageux pour les intervenants d'être en mesure de comparer les résultats aux deux instruments. Cela pourrait leur permettre d'identifier le niveau d'intervention nécessaire dans l'EIS à partir des résultats obtenus à l'ASQ-3 pour un même enfant. De cette façon, l'intervenant serait en mesure de débiter directement son intervention selon les étapes de développement déjà atteintes par l'enfant. Il est toutefois nécessaire de s'assurer que la distance par rapport aux scores seuils des deux instruments soit comparable. Il serait également possible d'identifier à partir d'un score à l'EIS les enfants qui ont un développement significativement inférieur à la moyenne.

Dans un premier temps, la comparaison entre les scores aux deux instruments montre que l'EIS est en mesure d'identifier avec un haut taux de réussite les enfants de l'échantillon utilisé présentant potentiellement un retard significatif dans les deux sphères concernées. En effet, la grande majorité des enfants montrant un développement typique à l'ASQ ont été considérés sans problématique par l'EIS. Il importe également d'investiguer les résultats obtenus aux différentes corrélations effectuées. En effet, les résultats montrent des corrélations très faibles, négatives et non significatives. Le petit échantillon utilisé ainsi que les sept résultats divergents entre l'EIS et l'ASQ pourraient expliquer ces résultats. De plus, il est possible de constater que les scores obtenus

à l'ASQ-3 varient de façon beaucoup plus importante que ceux obtenus à l'EIS. Cela fait en sorte que les deux variables impliquées dans chacune des corrélations ne varient pas de façon homogène, ce qui pourrait également contribuer au fait que les corrélations ne soient pas significatives. Un autre facteur pouvant avoir influencé les résultats, soit les données extrêmes, est observable dans les diagrammes de dispersions.

Il est intéressant de constater que les enfants ayant obtenu des résultats à l'EIS qui divergent de ceux à l'ASQ-3 se retrouvent majoritairement au niveau un de l'instrument. Bien que cela puisse être causé par le hasard, différentes hypothèses pourraient expliquer ces résultats. Premièrement, le nombre d'items dans le niveau un est plus élevé que dans le niveau deux, cela fait en sorte que les scores seuils sont également plus élevés. Toutefois, comme l'EIS n'a pas été conçu pour déterminer ou non la présence d'un retard de développement, il est possible que les items aient été développés dans le but d'être très sensibles aux progrès de l'enfant mais qu'ils manquent de spécificité. De plus, un plus grand nombre d'items augmente le risque d'erreur de la part de l'évaluateur. Les différents intervalles de scores seuils utilisés peuvent également jouer un rôle dans ces résultats. D'abord, il faut spécifier que l'EIS utilise des tranches d'âge de trois mois pour le niveau un. L'étude de Bricker *et al.*, (2008) montre que la spécificité pour la tranche d'âge dans laquelle se retrouvent trois des sept résultats divergents, soit de 34 à 36 mois, est de 63 %. Cela signifie qu'un enfant sur trois sera identifié comme présentant potentiellement un retard développemental par l'EIS alors qu'il n'en présente pas. Cela pourrait expliquer la forte concentration de résultats divergeant dans cette tranche d'âge.



### 3. LIMITES DU PROJET

Bon nombre de facteurs méthodologiques ont pu influencer les résultats. Pour la plupart, ces facteurs sont des répercussions directes des circonstances particulières ayant entouré la collecte de données du projet. Tout d'abord, l'échantillon utilisé pour les analyses est de 21 enfants, ce qui se trouve à la limite minimale pour une corrélation de Pearson (David, 1938). La cotation de l'EIS s'est effectuée uniquement avec les cotes « 0 » et « 2 » et un enfant qui aurait habituellement obtenu la cote « 1 » obtenait la cote « 0 ». Cette procédure est en concordance avec l'étude de Bricker *et al.* (2003) qui visait le développement des scores seuils pour la version anglophone de l'EIS. Toutefois, cette correction a eu pour effet de réduire la variance au sein des données et donc à homogénéiser l'échantillon. En effet, il est possible que certains enfants ayant obtenu des résultats divergeant avec cette façon de faire auraient obtenu un résultat différent si la cote « 1 » avait été utilisée. La formule de la corrélation de Pearson est très sensible à la variance des données puisqu'elle utilise la covariance, c'est-à-dire la façon dont deux variables varient ensemble, pour déterminer un coefficient (Schober, Boer et Schwarte, 2018). Finalement, les enfants recrutés ne présentent aucune condition clinique et leurs résultats à l'ASQ-3 sont très élevés. Cela fait en sorte que l'échantillon ne se distribue pas normalement, ce qui va à l'encontre d'un des postulats de la corrélation de Pearson qui affirme que pour l'échantillon utilisé doit suivre une courbe normale (Swank et Mullen, 2017). Cette situation a donc pu affaiblir la puissance statistique des corrélations.

#### 4. RECHERCHES FUTURES

Compte tenu du contexte particulier et des lacunes méthodologiques précédemment mentionnées, il serait pertinent de reproduire le projet en prenant en considération certaines modifications. Tout d'abord, la cotation de l'EIS devrait être complète, c'est-à-dire que non seulement les buts mais également les objectifs devraient être cotés en utilisant l'échelle à trois scores habituellement utilisée, soit « 0 », « 1 » ou « 2 ». De plus, un échantillon plus important devrait être recruté afin de maximiser la puissance statistique de la corrélation de Pearson. Cet échantillon devrait aussi contenir des enfants ayant différents niveaux développementaux afin d'assurer une normalité statistique de l'échantillon. De plus, il serait intéressant de s'intéresser aux différences observables entre les deux niveaux de l'EIS et plus spécifiquement aux différentes tranches d'âge utilisées. L'ensemble des domaines des deux outils pourraient également être comparés afin d'identifier si des tendances sont présentes pour certains domaines en particulier. Finalement, différentes contraintes ont fait en sorte que le présent projet n'a pas été en mesure d'évaluer la capacité de l'EIS à identifier un retard de développement significatif auprès d'enfants ayant été préalablement évalués et diagnostiqués. La prochaine étape serait donc d'effectuer la même étude auprès d'enfants présentant un RGD.

## CONCLUSION

Le présent mémoire a comme objectif de comparer les résultats obtenus dans deux domaines de développement par un instrument de dépistage et ceux obtenus par un instrument d'évaluation authentique du développement de l'enfant dans l'optique de déterminer si une comparaison entre les deux pourrait s'avérer possible. Il s'inscrit dans une tentative de réponse au manque d'instruments d'évaluation authentique du développement validés en français. Ce mémoire propose donc une piste de solution pour les recherches futures sur le sujet en démontrant que l'EIS, un instrument d'évaluation authentique, semble en mesure d'identifier adéquatement les enfants ne présentant pas de retard développemental connu dans les sphères cognitives et de la motricité fine. Il contribue également à démontrer l'importance de la complémentarité entre les instruments d'évaluation du développement standardisés et les instruments d'évaluation authentique. Bien que les résultats des corrélations s'avèrent négatifs et non significatifs, il est possible de constater que les résultats aux deux instruments sont cohérents pour la grande majorité des enfants de l'échantillon.

Ces résultats pourraient contribuer à la validation francophone de l'EIS en ce qui a trait à la détermination de l'éligibilité aux services pour les jeunes enfants présentant potentiellement un retard global de développement. Cela permettrait aux professionnels impliqués dans l'évaluation des jeunes enfants d'utiliser un instrument d'évaluation authentique n'étant pas affecté par l'ensemble des lacunes préalablement nommées concernant les instruments d'évaluation du développement standardisés pour corroborer une décision d'éligibilité avec justesse. Toutefois, les recherches futures devront se concentrer sur la réalisation de ce type d'étude à plus grande échelle,

dans un environnement francophone. Le développement de la littérature francophone concernant les instruments d'évaluation authentique du développement pourrait permettre aux intervenants directement concernés d'avoir sous la main davantage d'outils leur permettant d'effectuer des évaluations et des interventions efficaces et qui, au final, seront bénéfiques pour les jeunes enfants.

## RÉFÉRENCES BIBLIOGRAPHIQUES

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders*. Washington, DC: Author.

Bagnato, S. J., Goins, D. D., Pretti-Frontczak, K. et Neisworth, J. T. (2014). Authentic assessment as 'best practice' for early childhood intervention: National consumer social validity research. *Topics in Early Childhood Special Education*, 34(2), 116-127. doi: 10.1177/0271121414523652

Bagnato, S. J. et Ho, H. Y. (2006). High-stakes testing with preschool children: Violation of professional standards for evidence-based practice in early childhood intervention. *KEDI Journal of Educational Policy*, 3(1), 23-43.

Bagnato, S. J., Neisworth, J. T. et Pretti-Frontczak, K. (2010). *Linking authentic assessment & early childhood intervention best measure for best practices* (2<sup>e</sup> éd.). Baltimore, MD: Paul H. Brookes Publishing Co.

Baurain, C. et Nader-Grosbois, N. (2009). Évaluer la régulation émotionnelle, la résolution de problèmes socio-émotionnels et les compétences sociales d'enfants présentant une déficience intellectuelle: Études de cas. *Revue Francophone de la Déficience Intellectuelle*, 20, 123-147.

Bayley, N. (1969). *Bayley scales of infant and toddler development*. San Antonio, TX: Psychological Corporation.

Bigras, N., Tremblay, M. et Lemay, L. (2012). *Petite enfance, services de garde éducatifs et développement des enfants : état des connaissances*. Québec, QC: Presses de l'Université du Québec.

Bricker, D. (1993). *Assessment, Evaluation, and Programming System for Infants and Children*. Baltimore, MD : Paul H. Brookes Publishing.

Bricker, D. (2002). *Assessment evaluation and programming system for infants and children* (2<sup>e</sup> éd.). Baltimore, MD: Paul H. Brookes Publishing.

Bricker, D. (2006). *Programme d'évaluation, d'intervention et de suivi. Tome I : guide d'utilisation et tests : 0 à 6 ans*. Traduction et adaptation auprès d'une clientèle québécoise sous la direction de C. Dionne en collaboration avec C.-A. Tavares et C. Rivest. Montréal, QC: Chenelière McGraw Hill.

Bricker, D., Bailey, E. J. et Slentz, K. (1990). Reliability, validity, and utility of the evaluation and programming system: For infants and young children (EPS-I). *Journal of Early Intervention*, 14(2), 147-160. doi: 10.1177/105381519001400204

- Bricker, D., Clifford, J., Yovanoff, P., Pretti-Frontczak, K., Waddell, M., Allen, D. et Hoselton, R. (2008). Eligibility determination using a curriculum-based assessment: A further examination. *Journal of Early Intervention*, 31(1), 3-21. doi: 10.1177/1053815108324422
- Bricker, D. et Squires, J. (1999). *The ages and stages questionnaires: A parent-completed, child-monitoring system* (2e éd.). Baltimore, MD: Brookes.
- Bricker, D., Squires, J. et Clifford, J. (2010). Developmental Screening Measures: Stretching the Use of the ASQ for Other Assessment Purposes. *Infants and Young Children*, 23(1), 14-22.
- Bricker, D., Yovanoff, P., Capt, B. et Allen, D. (2003). Use of a curriculum-based measure to corroborate eligibility decisions. *Journal of Early Intervention*, 26(1), 20-30. doi: 10.1177/105381510302600102
- Dalton, E. M. et Brand, S. T. (2012). The Assessment of Young Children through the Lens of Universal Design for Learning (UDL). *Forum on Public Policy Online*, 2012(1).
- Danielsson, H., Henry, L., Messer, D. et Rönnberg, J. (2012). Strengths and weaknesses in executive functioning in children with intellectual disability. *Research in Developmental Disabilities*, 33(2), 600-607. doi: 10.1016/j.ridd.2011.11.004

- David, F. N. (1938). *Tables of the ordinates and probability integral of the distribution of the correlation coefficient in small samples*. Cambridge, UK: Cambridge University Press.
- Dennis, L. R., Rueter, J. A. et Simpson, C. G. (2013). Authentic Assessment: Establishing a Clear Foundation for Instructional Practices. *Preventing School Failure*, 57(4), 189-195.
- Division for Early Childhood. (2009). *Developmental delay as an eligibility category (Concept Paper)*. Missoula, MT: Author.
- Dhuey, E. et Lipscomb, S. (2010). Disabled or young? Relative age and special education diagnoses in schools. *Economics of Education Review*, 29(5), 857-872.
- Embretson, S. E. et Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum.
- Fortin, M.-F. (2010). *Fondements et étapes du processus de recherche : méthodes quantitatives et qualitatives* (2<sup>e</sup> éd.). Montréal. QC: Chenelière Éducation.
- Freeman, N. et Brown, M. (2008). An Authentic Approach to Assessing Pre-Kindergarten Programs: Redefining Readiness. *Childhood Education*, 84(5), 267.



- Gendreau, G. (2001). *Jeunes en difficulté et intervention psychoéducative*. Montréal, Québec : Éditions Sciences et Culture.
- Gao, X. et Grisham-Brown, J. (2011). The use of authentic assessment to report accountability data on young children's language, literacy and pre-math competency. *International Education Studies*, 4(2), 41-53. doi: <http://dx.doi.org/1005539/ies.v4n2p41>
- Gilbert, R., Logan, S., Moyer, V. A. et Elliott, E. J. (2001). Assessing diagnostic and screening tests. *The Western journal of medicine*, 174(6), 405-409. doi: 10.1136/ewjm.174.6.405
- Glaser, B. et Strauss, A. (1967). *The discovery of grounded theory : Strategies for qualitative research*. Chicago, IL: Aldine.
- Grisham-Brown, J., Hallam, R. A. et Pretti-Frontczak, K. (2008). Preparing head start personnel to use a curriculum-based assessment: An innovative practice in the 'age of accountability'. *Journal of Early Intervention*, 30(4), 271-281. doi: 10.1177/1053815108320689
- Haley, S. M., Coster, W., Ludlow, L. H., Haltiwanger J. T. et Andrellos, P. J. (1992). *Pediatric Evaluation of Disability Inventory: Development, Standardization, and Administration Manual*. Boston, MA: New England Medical Center Hospital/Trustees of Boston University.

Hallam, R. A., Lyons, A. N., Pretti-Frontczak, K. et Grisham-Brown, J. (2014). Comparing Apples and Oranges: The Mismeasurement of Young Children through the Mismatch of Assessment Purpose and the Interpretation of Results. *Topics in Early Childhood Special Education*, 34(2), 106-115.

Harrison, P. L. et Oakland, T. (2003). *Adaptive Behavior Assessment System-Second Edition*. San Antonio, TX: The Psychological Corporation.

High/Scope Educational Research Foundation. (2003). *Preschool Child Observation Record—Second Edition*. Ypsilanti, MI: High/Scope Press.

Hresko, W. P., Miguel, S. A., Sherbenou, R. J. et Burton, S. D. (1994). *Developmental Observation Checklist System*. Austin, TX: PRO-ED.

Individuals with Disabilities Education Act, 20 U.S.C. § 1400 (2004)

Institut national de santé publique du Québec. (2010). *Avis scientifique sur le choix d'un outil de mesure du développement des enfants de 0 à 5 ans dans le cadre des Services intégrés en périnatalité et pour la petite enfance*. Repéré à <http://collections.banq.qc.ca/ark:/52327/1990262> Accès au document via la BAnQ

- Johnson, B. (2016). Early Intervention in Children with Developmental Disabilities. *BMH Medical Journal*, 3(1), 1-4.
- Johnson, J. J. et Bricker, D. (Éds). (sous presse). *Evaluation, Assessment, and Programing System* (3<sup>e</sup> éd.). Baltimore, MD: Paul H. Brookes Publishing Co.
- Keilty, B., LaRocco, D. J. et Casell, F. B. (2009). Early Interventionists' Reports of Authentic Assessment Methods Through Focus Group Research. *Topics in Early Childhood Special Education*, 28(4), 244-256.
- Kim, S. W., Jeon, H. R., Park, E. J., Chung, H. J. et Song, J. E. (2014). The Differences in Clinical Aspect Between Specific Language Impairment and Global Developmental Delay. *Ann Rehabil Med*, 38(6), 752-758.
- Kim, D.-H., Lambert, R. G. et Burts, D. C. (2013). Evidence of the validity of Teaching Strategies GOLD® assessment tool for English language learners and children with disabilities. *Early Education and Development*, 24(4), 574-595. doi: 10.1080/10409289.2012.701500
- Knobloch, H., Stevens, F et Malone, A. F. (1987). *Manual of developmental diagnosis: The administration and interpretation of the revised Gesell and Amatruda developmental and neurological examination*. Houston: Developmental Evaluation Materials, Inc.

- Koo, T. K. et Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of chiropractic medicine*, 15(2), 155-163. doi: 10.1016/j.jcm.2016.02.012
- Lazaro, S. L. (2017). The importance of authentic assessments in eligibility determination for infants and toddlers. *Journal of Early Intervention*, 39(2), 88-105. doi: 10.1177/1053815116689061
- Lee, D. D., Bagnato, S. J. et Pretti-Frontczak, K. (2015). Utility and validity of authentic assessments and conventional tests for international early childhood intervention purposes: Evidence from U.S. national social validity research. *Journal of Intellectual Disability: Diagnosis and Treatment*, 3(4), 164–176.
- Linacre, J. M. (2017). *Winsteps* (Version 4.0.1) [Computer software]. Chicago: Mesa Press.
- Linder, T.W. (1993). *Transdisciplinary play-based assessment: A functional approach to working with young children*. Baltimore, MD: Brookes.
- Macy, M. et Bagnato, S. J. (2010). Keeping It "R-E-A-L" with Authentic Assessment. *NHSA Dialog*, 13(1), 1-20.

Macy, M. et Bagnato, S. J. (2013). The authentic alternative for assessment in early childhood intervention. Dans D. H. Saklofske, C. R. Reynolds, V. L. Schwann, D. H. Saklofske, C. R. Reynolds et V. L. Schwann (dir.), *The Oxford handbook of child psychological assessment*. (p. 671-682). New York, NY, US: Oxford University Press.

Macy, M., Bagnato, S. J. et Gallen, R. (2016). Authentic Assessment. *Zero to Three*, 37(1), 37-43.

Macy, M., Bagnato, S. J., Macy, R. S. et Salaway, J. (2015). Conventional tests and testing for early intervention eligibility: Is there an evidence base? *Infants & Young Children*, 28(2), 182-204. doi: 10.1097/iy.0000000000000032

Macy, M. G., Bricker, D. D. et Squires, J. K. (2005). Validity and Reliability of a Curriculum-Based Assessment Approach to Determine Eligibility for Part C Services. *Journal of Early Intervention*, 28(1), 1-16.

Macy, M. et Hoyt-Gonzales, K. (2007). A Linked System Approach to Early Childhood Special Education Eligibility Assessment. *TEACHING Exceptional Children*, 39(3), 40-44.

Macy, M., Thorndike-Christ, T. et Lin, Y.-C. (2010). Parental reports of perceived assessment utility: A comparison of authentic and conventional approaches. *Infants & Young Children*, 23(4), 286-302. <http://dx.doi.org/10.1097/IYC.0b013e3181f1ec1b>

Masters, G. (1988). The analysis of partial credit scoring. *Applied Measurement in Education*, 1, 279-297.

McLean, M. (2005). Using Curriculum-Based Assessment to Determine Eligibility: Time for a Paradigm Shift? *Journal of Early Intervention*, 28(1), 23-27.

Meisels, S. J., Wen, X. et Beachy-Quick, K. (2010). Authentic assessment for infants and toddlers: Exploring the reliability and validity of the Ounce Scale. *Applied Developmental Science*, 14(2), 55-71. doi: 10.1080/10888691003697911

Ministère de la famille. *Directive concernant l'allocation pour l'intégration d'un enfant handicapé*. Québec, QC : Gouvernement du Québec.

Ministère de la santé et des services sociaux. (2017). Québec, QC: Gouvernement du Québec.

Moreno, A. J. et Klute, M. M. (2011). Infant–toddler teachers can successfully employ authentic assessment: The Learning Through Relating system. *Early Childhood Research Quarterly*, 26(4), 484-496. doi: 10.1016/j.ecresq.2011.02.005

- National Association for Education of Young Children. et National Association of Early Childhood Specialists in State Departments of Education. (2003). *Early childhood curriculum assessment and program evaluation*. National Association for Education of Young Children. Repéré à <http://www.naeyc.org/files/naeyc/file/positions/pscape.pdf>
- Neisworth, J. T. et Bagnato, S. J. (2004). The MisMeasure of Young Children. *Infants & Young Children: An Interdisciplinary Journal of Early Childhood Intervention*, 17(3), 198-212.
- Newborg, J., Stock, J.R., Wnek, L., Guidubaldi, J. et Svinicki, J. (1984). *Battelle Developmental Inventory*. Itasca, IL: Riverside.
- Newborg, J., Stock, J., Wnek, L., Guidubaldi, J. et Svinicki, J. (1988). *Battelle Developmental Inventory: Examiner's manual*. Dallas, TX: DLM/Teaching Resource.
- Newborg, J., Stock, J., Wnek, L., Guidubaldi, J. et Svinicki, J. (2004). *Battelle Developmental Inventory* (2e éd.). Chicago, IL: Riverside Publishing.
- Noh, J. (2005). *Examining the psychometric properties of the second edition of the Assessment, evaluation, and programming system for three to six years : AEPS test 2nd Edition (3-6)* (Thèse de doctorat inédite). University of Oregon. Repéré à <http://search.ebscohost.com/login.aspx?direct=true&db=psych&AN=2006-99013-33&site=ehost-live>

Notari-Syverson, A. et Losardo, A. (2004). What Assessment Means to Early Childhood Educators. *Exchange*, (159), 72-75.

Office of Special Education and Rehabilitative Services. (2002). *A new era: Revitalizing special education for children and their families*. Washington, DC: Author.

Ordre des psychoéducateurs et psychoéducatrices du Québec [OPPQ] (2014). *L'évaluation psychoéducative de la personne en difficulté d'adaptation. Lignes directrices*. Montréal, QC: Auteur.

Ordre des psychoéducateurs et psychoéducatrices du Québec [OPPQ] (2013). *Évaluation du retard de développement : lignes directrices*. Montréal, QC: Ordre des psychoéducateurs et psychoéducatrices du Québec.

Parks, S. (2006). *Inside HELP: Administration and reference manual for using the Hawaii Early Learning Profile as a birth to three, curriculum based assessment*. Palo Alto, CA: VORT Corporation.

Parks, S., Furono, S., O'Reilly, K., Inatsuka, T., Hoska, C. M. et Zeisloft-Falbey, B. (1994). *Hawaii early learning profile (HELP)*. Palo Alto, CA: Vort.



- Pretti-Frontczak, K. L. (2002). Using curriculum-based measures to promote a linked system approach. *Assessment for Effective Intervention*, 27(4), 15–21.
- Rousseau, M., Dionne, C. et Paquet, A. (2012). Intervention précoce et instruments d'évaluation des jeunes enfants ayant un trouble du spectre autistique. *Revue francophone de la déficience intellectuelle*, 23, 93-103.
- Scherzer, A. L., Chhagan, M., Kauchali, S. et Susser, E. (2012). Global perspective on early diagnosis and intervention for children with developmental delays and disabilities. *Developmental medicine and child neurology*, 54(12). doi: 10.1111/j.1469-8749.2012.04348.x
- Schober, P., Boer, C. et Schwarte, L. A. (2018). Correlation Coefficients: Appropriate Use and Interpretation. *Anesthesia and analgesia*, 126(5), 1763-1768.
- Slentz, K. (2008). *What AEPS® users should know about reliability, validity, and utility*. Batlimore, MD: Paul H. Brookes Publishing Co.
- Statistique Canada. (2001). *Enquête sur la participation et les limitations d'activités*. Repéré à <http://www.statcan.gc.ca/pub/89-577-x/t/4065018-fra.htm>

- Shevell, M. (2008). Global Developmental Delay and Mental Retardation or Intellectual Disability: Conceptualization, Evaluation, and Etiology. *The Pediatric Clinics of North America*, 55, 1071-1084. doi: 10.1016/j.pcl.2008.07.010
- Shevell, M., Ashwal, S., Donley, D., Flint, J., Gingold, M., Hirtz, D., . . . Sheth, R. D. (2003). Practice parameter: Evaluation of the child with global developmental delay: Report of the Quality Standards Subcommittee of the American Academy of Neurology and The Practice Committee of the Child Neurology Society. *Neurology*, 60(3), 367-385. doi: 10.1212/01.WNL.0000031431.81555.16
- Shackelford, J. (2006). *State and jurisdictional eligibility definitions for infants and toddlers with disabilities under IDEA* (NECTAC Notes No. 21). Chapel Hill: The University of North Carolina, FPG Child Development Institute, National Early Childhood Technical Assistance Center.
- Sparrow, S. S., Balla, D. A. et Cicchetti, D. V. (1984). *Vineland adaptive behavior scales: Interview edition survey form manual*. (2<sup>e</sup> éd.). Circle Pines, MN: American Guidance Service.
- Squires, J. et Bricker, D. D. (2009). *Ages & stages questionnaires : a parent-completed child monitoring system* (3<sup>e</sup> éd.). Baltimore: Paul H. Brooks Pub. Co.

Squires, J., Twombly, E., Bricker, D. et Potter, L. (2009). *ASQ-3. User's Guide* (3e éd.). Baltimore: Brookes Publishing Co.

Srour, M., Mazer, B. et Shevell, M. (2006). Analysis of clinical features predicting etiologic yield in the assessment of global developmental delay. *Pediatrics*, 118(1), 139-145.

Taber, K. S. (2018). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research in Science Education*, 48(6), 1273-1296. doi: 10.1007/s11165-016-9602-2

Tardif, G., Bérubé, A., Coutu, S., Dubeau, D., Jacques, C. et Lemay, L. (2015). *Synthèse des connaissances empiriques et cliniques : proposition d'outils permettant d'adapter ou de construire une mesure populationnelle du développement de l'enfant de 2 ans*. Québec : Fonds de recherche du Québec - Société et culture (FRQSC).

Waddell, M., Pretti-Frontczak, K., Johnson, J. et Bricker, D. (2007). *Using AEPS® to determine eligibility for IDEA services*. Baltimore: Paul H. Brookes Publishing Co